National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

University of Alberta

Conic Model Methods with a Trust Region Approach and Calmness of Optimization Problems

BY

Simon DI

A Thesis

Submitted to the Faculty of Graduate Studies and Research

in Partial Fulfillment of the Requirements for the Degree

of Doctor of Philosophy

Department of Mathematics

Edmonton, Alberta

Fall 1992

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-77231-X

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **Simon DI**

TITLE OF THESIS:

Conic Model Methods with a Trust Region Approach and Calmness of Optimization Problems

DEGREE: **Doctor of Philosophy**

YEAR THIS DEGREE GRANTED: **1992**

Permission is hereby granted to the UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

(Signed) Simon Di

Permanent Address:
Department of Mathematics
University of Alberta
Edmonton, Alberta
Canada T6G 2G1

Date: Aug. 31st, 1992

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled

Conic Model Methods with a Trust Region Approach and Calmness of Optimization Problems

submitted by **Simon DI** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Optimization Theory and Application**.

Dr. R. Poliquin, supervisor

Dr. J.W. Macki

Dr. J.S. Muldowney

Dr. B. Joe

Dr. Y.S. Wong

Dr. D. Ward, external examiner

Date: Aug. 25, 1992

To:

David

Julie

FengYa

Abstract

In this thesis, we make a contribution to two aspects of optimization theory and applications. A new numerical method for solving unconstrained minimization problems is proposed, and explicit expressions of various kinds of normal cones are established and used to develop sufficient conditions for calmness.

In **Part I**, which includes **Chapter 2** to **Chapter 5**, conic model methods with trust region approach are proposed. We discuss their motivation and theoretical background, which involve necessary and sufficient conditions for solutions to the model trust region subproblems. A specific computer program shows that the class of methods has very satisfactory numerical performance for most of the standard test functions. The convergence properties proved in the thesis show that a typical algorithm has Q-superlinear rate of convergence. All the topics suggest that the methods we propose are very competitive for solving unconstrained minimization problems.

In **Part II**, which includes **Chapter 6** to **Chapter 8**, some sufficient conditions for calmness are established. We begin this part by recalling that "*almost all problems are calm*" for inequality-constrained optimization problems, and we consider whether the same assertion is true for optimization problems that incorporate equality constraints and an abstract constraint. We provide counterexamples for the open problem and develop some sufficient conditions for calmness based on the relationships between the normal cone of the equality constraint set and the generalized gradients of the constraint functions. These relationships enrich the theory of generalized gradients (Jacobian) and various kind of normal cones provide explicit expressions for the normal cones of certain nonsmooth manifolds in finite dimensional space.

Acknowledgements

I'd like to thank my supervisor Dr. René Poliquin, with deepest appreciation, for his invaluable guidance and friendly support.

I also wish to thank Dr. Jack Macki, Dr. James Muldowney, Dr. Yau Shu Wong for their constant advice and interest in my research.

I am grateful to Dr. Doug Ward for his careful proofreading of this thesis and his thoughtful suggestions to correct some of the errors in the preliminary drafts of the thesis.

I thank the Department of Mathematics for providing congenial research environment. I am grateful to Dr. René Poliquin and the University of Alberta for their financial support.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION 1

1.1 Problems and Definitions 2

1.2 Numerical Solutions and Conic Model Methods with Trust Region Approach 6

1.3 Outline of the Thesis 8

CHAPTER 2. GENERAL SURVEY OF CURRENT ALGORITHMS FOR SOLVING UNCONSTRAINED MINIMIZATION PROBLEMS 12

2.1 Review of Classical Methods 12

2.2 Line Searches and the Trust Region Approach 16

CHAPTER 3. MOTIVATIONS OF THE TRUST REGION APPROACH FOR CONIC MODELS 20

3.1 Conic Models 21

3.2 Collinear Scaling Conic Models 25

3.3 The Trust Region Approach for Conic Models and its Subproblems 29

CHAPTER 4. NUMERICAL IMPLEMENTATION AND COMPUTATIONAL RESULTS 36

4.1 Numerical Implementation 36

4.2 Testing and Computational Results 40

CHAPTER 5. CONVERGENCE RESULTS AND CONCLUSIONS 48

5.1 Convergence Properties 48

5.2 Conclusions 58

CHAPTER 6. CALMNESS AND COUNTEREXAMPLES 60

6.1 Calmness and Constraint Qualifications 60

6.2 Calmness and its other Properties 62

6.3 Counterexamples 66

CHAPTER 7. TANGENT CONES AND

NORMAL CONES TO THE CONSTRAINT SET 70

7.1 Definitions and Review of the Results 70

7.2 Results for One Equality Constraint 75

7.3 The Clarke Normal Cone to the Equality Constraint Set 79

7.4 The Contingent Cone and the Strict Normal Cone

at a Differentiable Point 84

7.5 More Results about the Clarke Normal Cone 88

CHAPTER 8. CALMNESS AND THE GENERALIZED JACOBIAN 92

8.1 Calmness and the Normal Cone to the Constraint Set 92

8.2 Calmness Under More Favorable Conditions 95

8.3 Conclusions 99

APPENDIX 101

REFERENCES 104

CHAPTER 1

INTRODUCTION

In the classical sense, optimization can be defined as the art of obtaining best policies to satisfy certain objectives, while satisfying fixed requirements. Recent advances in numerical mathematics, operations research, and computer technology enable many complex industrial problems in engineering, economics, and management science to be optimized successfully by the application of logic and systematic techniques. The development of new and increasingly powerful optimization techniques is proliferating rapidly and the theory of optimization is already an established science. However, more work is needed to expand the applications of optimization techniques in detailed complex industrial problems, to develop highly efficient numerical methods to keep pace with the daily changing digital computer techniques, as well as to enrich the theory.

To study a given practical problem it is necessary to form a mathematical cause-and-effect relationship — commonly known as a *mathematical model*. Such a model must describe accurately the laws of economics and human behavior, as well as the natural relationships among all the factors involved in the problem. In certain fields of practical applications, practical applications of optimization theory are detained and doubted sometimes, because it is usually impossible for a mathematical model to match completely the practical system. However, it is necessary and sufficient to model an entire system in the broadest, most accurate possible sense to get a practical problem solved.

Once a mathematical model is formulated, a computer is usually used to obtain its solution. The fast development of modern computers has resulted in higher

speeds of numerical computing, huge amounts of memory and quick exchanges of input and output information. But how to make use of the advances in hardware and software technology in order to improve our mathematical model, to have corresponding faster numerical methods (such as using distributive computing) is always a challenging problem.

From the theoretical point of view, it is most important to discuss the necessary and sufficient optimality conditions for different kinds of optimization problems. The necessary and sufficient conditions provide criteria in establishing numerical methods. For the smooth problems, such as linear and nonlinear programming problems, the theory is pretty mature now. However there is still much research to be done in nonsmooth optimization theory.

In this thesis, we attempt to make contributions to the two issues; namely, the construction of a new algorithm for solving unconstrained nonlinear optimization problems; and a discussion of calmness, which serves as a constraint qualification in establishing necessary optimality conditions.

1.1 Problems and Definitions

Mathematical programming is a branch of optimization theory (see e.g. [Avri76]). In mathematical programming, the problem is to minimize (or maximize) a single objective function subject to a number of constraints including equality constraints, inequality constraints and an abstract constraint, i.e. we deal with prob-

lem $\mathcal{P}$:

$$\begin{cases} min\ f(x) \\ s.t.\ \ g_i(x) \le 0 \ \ i = 1,2,\dots,s, \\ \qquad h_j(x) = 0 \ \ j = 1,2,\dots,m, \\ \qquad x \in C, \end{cases} \tag{1.1.1}$$

where $f, g_i, h_j : \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz continuous for $i = 1,2,\dots,s$, $j = 1,2,\dots,m$ and C is a given closed subset of $\mathbb{R}^n$. Let $g = (g_1, g_2, \dots, g_s)$, and $h = (h_1, h_2, \dots, h_m)$; then $g : \mathbb{R}^n \to \mathbb{R}^s$, and $h : \mathbb{R}^n \to \mathbb{R}^m$. Throughout this thesis we use $|\cdot|$ or $\|\cdot\|$ to denote the Euclidean norm.

DEFINITION 1.1.1. *A point $x^* \in \mathbb{R}^n$ is called a local minimizer for problem $\mathcal{P}$ if x^* satisfies all the constraints of $\mathcal{P}$ and there is a neighborhood $\mathcal{N}$ of x^* such that $f(x) \ge f(x^*)$ for all $x \in \mathcal{N}$ satisfying the constraints of $\mathcal{P}$.*

The most important necessary condition for x^* to be a local minimizer of problem $\mathcal{P}$ is the *Lagrange multiplier rule* which was proved for nonsmooth functions by F.H. Clarke (see [Clar76]) by using the *generalized gradient* defined by himself.

DEFINITION 1.1.2. *Given a function $f : \mathbb{R}^n \to \mathbb{R}$, the generalized directional derivative of f, evaluated at x in the direction v, is defined by*

$$f^\circ(x;v) := \limsup_{y \to x,\, t \downarrow 0} \frac{f(y+tv) - f(y)}{t}. \tag{1.1.2}$$

For a Lipschitz function f, the function: $v \to f^\circ(x;v)$ is always finite, positive homogeneous and subadditive. The function $f^\circ(x;v)$ is always upper semicontinuous as a function of (x,v) and $f^\circ(x;-v) = (-f)^\circ(x;v)$ (see [Clar83]).

DEFINITION 1.1.3. *The set*

$$\partial f(x) := \{\, \xi \in \mathbb{R}^n : f^\circ(x;v) \geq \langle v, \xi \rangle \quad \textit{for all} \quad v \in \mathbb{R}^n \,\} \tag{1.1.3}$$

is called the generalized gradient of f *at* x .

The generalized gradient of a locally Lipschitz function is a nonempty compact subset of $\mathbb{R}^n$. According to Rademacher's theorem, a Lipschitz function $f(x)$ is differentiable almost everywhere (in the sense of *Lebesgue measure*). If we denote by Ω_f the set of points at which f fails to be differentiable, then the generalized gradient has an equivalent expression (see [Clar83]):

$$\partial f(x) := co\{\, \lim \nabla f(x_i) : x_i \to x, x_i \notin S \cup \Omega_f \,\}, \tag{1.1.4}$$

for any given set S of *Lebesgue measure* 0 in $\mathbb{R}^n$, where " *co* " denotes the convex hull.

The Lagrange multiplier rule, which is in §6.1, can be viewed as degenerate when the multiplier associated with the objective function is zero (because the function being minimized is not involved). Various supplementary conditions called *constraint qualifications* have been developed to avoid this awkwardness. The most widely used one is the *Mangasarian-Fromowitz* condition, and *calmness* is another one, as *Proposition 6.1.5* shows.

It is often desirable to have information about the *sensitivity* of the optimal solutions and optimal value to various perturbations. For example, in data-fitting problems, it may be important to know the effect on the solution of errors in the

data. In this sense, we have to consider the following problem $\mathcal{P}(p,q)$:

$$\begin{cases} min\ f(x) \\ s.t.\ \ g_i(x) \le p_i \ \ i = 1,2,\dots,s, \\ \qquad h_j(x) = q_j \ \ j = 1,2,\dots,m, \\ \qquad x \in C, \end{cases} \tag{1.1.5}$$

DEFINITION 1.1.4. *Let x be a local solution to $\mathcal{P}$. The problem $\mathcal{P}$ is calm at x provided that there exist positive ϵ and M such that for all (p,q) in $\epsilon\mathcal{B}$, and for all x' in $x+\epsilon\mathcal{B}$, feasible for $\mathcal{P}(p,q)$, one has*

$$f(x') - f(x) + M|(p,q)| \ge 0, \tag{1.1.6}$$

where $\mathcal{B}$ denotes the unit sphere in the corresponding finite dimensional space. The constants M and ϵ are called the modulus and radius of calmness for $\mathcal{P}$ at x.

From *Definition 1.1.4*, if the problem is calm at its solution, then the optimal value for a new problem with the same objective function and perturbed constraints will not be too much smaller than the original optimal value. From *Theorem 6.2.4*, calmness also ensures that some penalty methods are exact (exact penalization), which provides theoretical support for penalty methods to solve constrained nonlinear minimization problems by solving sequential unconstrained minimization problems.

The notion of calmness originated in the study of convex programs (see [Fiac83]). In **Part II**, which includes **Chapter 6** to **Chapter 8**, we discuss calmness for equality constrained nonlinear problem, and provide some sufficient conditions for calmness. The sufficient conditions are based on the relationships between the normal cones of the constraint set and the generalized gradients of

the equality constraint functions developed in **Chapter 7** of this thesis.

1.2 Numerical Solutions and Conic Model Methods with Trust Region Approach

In practical applications, looking for numerical solutions to mathematical programming problems is more important than discussing necessary and sufficient conditions for optimality. To find a numerical solution, it is necessary to use a computer program either from a numerical software library or developed by the problem analyst. The following characteristics of nonlinear problems are important in numerical computing. The first is the size of the problem. Most of the problems arising from practice have variables from 2 to 20 . The state of the art is such that we hope to solve most of the small problem, say those with from 2 to 15 variables. Intermediate problems in this field are those with from 15 to 50 variables; current algorithms can solve some of these. Problems with over 50 variables are considered to be large in this field (see [DeSc83]). Large problems are usually very difficult to solve unless they are only mildly nonlinear or there is a good starting guess. Therefore improving the efficiency of the computer algorithms for solving nonlinear minimization problems is a central issue in practical applications.

The second issue is the availability of derivatives. Frequently we are dealing with problems where the nonlinear objective function is the result of computer simulation or is given by a long and messy algebraic formula, and so the analytic derivatives are not readily available though the function is smooth. Therefore we have to use a computer algorithm to get the derivatives, for example by using finite difference methods. Except in very rare cases, we can approximate the analytic

derivatives and still have satisfactory convergence of the algorithm. Therefore all the algorithms can be carried out by using approximate derivatives as well as analytic derivatives.

The third issue concerns the fast development of hardware technology, including the availability of fast storage, fast central processing speed, huge amounts of random access memory and parallel central processing features. How to make full use of this advanced technology is an important topic.

Nonlinear minimization problems are divided into two classes: constrained and unconstrained. A constrained problem is usually solved by solving sequential unconstrained problems which approximate the original problem by using *penalty* or *barrier* methods. The penalty approach to constrained minimization is generally attributed to Courant [Cour62], and the barrier method is due to Carroll [Carr61], Fiacco and McCormick [FiMc68].

Since 1985, a number of methods for solving nonsmooth optimization problems have been developed. Among these methods, the bundle methods and their variations are considered to be most efficient and competitive (see [Zowe85], [Kiw85a], [Kiw85b], [ScZo88], [Schr89], [Kiw89a], [Kiw89b], [Kiw90] and [ScZo90]). These methods are quite different from the smooth minimization solvers in their appearance, but basically they use all the essential ideas of unconstrained minimization methods: including approximating the objective functions by some kinds of model functions (usually piecewise linear functions), finding the descent directions, and using line searches or the trust region approach.

In conclusion, methods for solving unconstrained nonlinear smooth minimization problems are essential to solve all kinds of nonlinear minimization problems. The development of such methods is still an important part of numerical optimiza-

tion. We propose a new class of methods of this kind in **Part I**, which includes **Chapter 2** to **Chapter 5**. Like any typical method for solving optimization problems, at each step of the method, we use a model function to approximate the objective function. By finding the minimum of the model function, we get a better estimate of the solution to the problem. In this thesis, we propose a conic model function method with a trust region approach. Suppose $f(x)$ is the objective function. We use a conic model function $\psi_k(s)$ to approximate $f(x_k+s)$ at the k-th iteration. There are two kinds of conic model functions. The first kind is a class of explicit conic functions:

$$\psi_k(s) = f_k + \frac{g_k^T s}{(1-a_k^T s)} + \frac{1}{2}\,\frac{s^T A_k s}{(1-a_k^T s)^2}, \tag{3.3.1}$$

proposed in [Davi80]. The second kind is a class of collinear scaling conic functions:

$$\begin{cases} \psi_k(s) &= f_k + g_k^T J_k w + \frac{1}{2} w^T B_k w, \\ s &= J_k w/(1+h_k^T w), \end{cases} \tag{3.3.2}$$

proposed in [Sore80]. Their trust region subproblem is

$$\min\{\psi_k(s) \mid \|D_k s\| \le \Delta_k\}, \tag{3.3.3}$$

where D_k is a scaling matrix, Δ_k is the trust region radius.

We present a necessary and sufficient condition for the solution to problem (3.3.3) for both kinds of conic functions. This solution plays an important role in the trust region method for the conic models, both theoretically and numerically. The motivations, theoretical background, numerical performance and convergence results are the main topics in **Part I** of this thesis.

1.3 Outline of the Thesis

This thesis is divided into two parts. **Part I**, which is from **Chapter 2** to **Chapter 5**, is concerned with the development of a class of unconstrained minimization methods. **Part II**, which is from **Chapter 6** to **Chapter 8**, establishes sufficient conditions for calmness and explicit expressions for some normal cones and tangent cones.

In **Chapter 2**, we present a general survey of the current algorithms for solving unconstrained minimization problems, including line searches and the trust region approach. **Chapter 2** summarizes all the unconstrained numerical optimization methods as methods equivalent to using sequential model approximations. The model functions are all quadratic functions.

Chapter 3 presents the motivations of the trust region approach for conic models. It takes a deep look at conic functions, which were proposed as approximation functions for solving unconstrained minimization in [Davi80]. The establishment of a collinear scaling conic model, which was done in [Sore80], is also discussed in the chapter. The core of the theory supporting the methods is the proof of the necessary and sufficient condition for the solutions to the trust region subproblems, which is done in the last part of the chapter.

Chapter 4 includes specifying the newly proposed algorithm, discussing the computer implementation and presenting the numerical performance on the standard test functions proposed in [MoGa81]. Some numerical features such as convergence criteria, scaling and testing are discussed, and the comparison of the method with one of the most advanced methods of its kind are presented in two tables.

In **Chapter 5**, we prove the convergence properties of a specific algorithm developed in **Chapter 4** under standard conditions on the objective function

and discuss some of the restrictions in the algorithm. The Q-superlinear rate of convergence proved in this chapter shows that the algorithm is very competitive with all the methods of its kind.

In **Chapter 6**, we review some definitions and applications of calmness, including its property as a constraint qualification, its equivalence to the exactness of certain penalty functions and an assertion that "*almost all problems are calm*" for inequality-constrained problems. For equality-constrained problems with an abstract constraint, this assertion does not hold because counterexamples are provided in **Chapter 6**. However, the question is still open for Lipschitz optimization problem incorporating only equality constraints.

In **Chapter 7**, we prove stronger results for the contingent cone to sets determined by equality and inequality constraints; these results require only differentiability at the given point instead of strict differentiability or differentiability in a neighborhood of the point. These kinds of relationships about the constraint set enrich the applications of the generalized gradient (Jacobian) and can be applied to characterize various kinds of the tangent cones and normal cones to the constraint set. An application to proximal normals of such sets at differentiable points of the constraint functions is presented. Another one to the Clarke normal cone of such sets is also presented under one more condition.

A relationship established in **Chapter 7** is used to prove a sufficient condition for calmness in **Chapter 8**. The results improve the current results available in the literature. We also provide two special cases of equality constraint problems, where the assertion "*almost all problems are calm*" is verified.

In the **Appendix**, we prove the intermediate value theorem in the general case, which is used to prove a theorem in **Chapter 7** via Brouwer degree.

PART I

A TRUST REGION METHOD FOR CONIC MODELS FOR SOLVING UNCONSTRAINED OPTIMIZATION PROBLEMS

CHAPTER 2

GENERAL SURVEY OF CURRENT ALGORITHMS FOR SOLVING UNCONSTRAINED OPTIMIZATION PROBLEMS

Throughout this part, we consider numerical methods for solving the unconstrained optimization problem:

$$\min f(x) \quad \text{for} \quad x \in \mathbb{R}^n, \tag{2.0.1}$$

for a function $f \in \mathcal{C}^2$, where $\mathcal{C}^k$ means the class of functions with continuous k-th derivative for a given integer k. We start by reviewing the classical methods for solving the problem and discuss their computer implementations and their good and bad features.

2.1 Review of Classical Methods

The steepest descent direction method is one of the earliest methods for solving multivariable minimization problems and dates back to Cauchy in 1847 [Cauc1847].

ALGORITHM 2.1.1. *(Steepest descent direction method) Given initial point* $x_0 \in \mathbb{R}^n$, *for* $k = 0, 1, 2, \cdots$, *until "convergence",*

1. *solve* $p_k = -\nabla f(x_k)$;
2. $x_{k+1} := x_k + p_k$.

The steepest descent direction method can be viewed as using a quadratic function to approximate the objective function at the iterative point by using the unit matrix as the Hessian of the quadratic function. The method has linear

rate of convergence, see [KaAk64] and [Luen73]. The convergence criteria will be discussed in §4.2.

ALGORITHM 2.1.2. *(Newton's method) Given initial point* $x_0 \in \mathbb{R}^n$, *for* $k = 0, 1, 2, \cdots$, *until "convergence",*

1. *calculate* $\nabla f(x_k)$, $\nabla^2 f(x_k)$ *and solve* $\nabla^2 f(x_k) p_k = -\nabla f(x_k)$;
2. $x_{k+1} := x_k + p_k$.

Newton's method can be viewed as using a quadratic function, or equivalently as using a second order Taylor's expansion, to approximate the objective function at the iterative point. The method has Q-quadratic rate of convergence. Because of the fast rate of convergence, Newton's method converges to a local solution very quickly when the initial guess is good enough and the Hessian at the local minimum is positive definite. But the method is not globally convergent and the price to pay for the fast convergence is to calculate the Hessian of the objective function at each iterative point, which is usually very costly.

A few modifications have been made to increase the efficiency of the above methods, e.g. using finite differences to approximate the gradient and the Hessian when the analytic gradient and Hessian are not readily available or difficult to evaluate, keeping the Hessian $\nabla^2 f(x_k)$ in factored form by using the *Cholesky factorization* given in [GMWr81]; and using $H_k = \nabla^2 f(x_k) + \nu_k I$ for some $\nu_k \geq 0$ instead of using $\nabla^2 f(x_k)$ in order to avoid the singular situation of the Hessian.

ALGORITHM 2.1.3. *(Conjugate gradient method) Given initial point* $x_0 \in \mathbb{R}^n$, *calculate* $\nabla f(x_0)$, *choose* $z_0 \in \mathbb{R}^n$ *s.t.* $z_0^T \nabla f(x_0) < 0$. *For* $k = 0, 1, 2, \cdots$, *until "convergence",*

1. $x_{k+1} := x_k + \lambda_k z_k$, *where* λ_k *minimizes* $f(x_k + \lambda z_k)$ *along the ray* $\{ x_k + \lambda z_k : \lambda \geq 0 \}$.

2. compute $\nabla f(x_{k+1})$, *calculate* $z_{k+1} := -\nabla f(x_{k+1}) + \beta_k z_k$, *where* $\beta_k := |\nabla f(x_{k+1})|^2 / |\nabla f(x_k)|^2$.

The method generates a sequence of directions $z_0, \cdots, z_k$ for which z_k is a linear combination of the preceding directions $z_0, \cdots, z_{k-1}$, and the current steepest descent direction $-\nabla f(x_k)$. If $f(x)$ is a quadratic function, i.e. $f(x) = \frac{1}{2} x^T A x + b^T x + c$ with A symmetric and positive definite, then the method minimizes f in at most n steps of iterations, and the direction vectors generated are A-orthogonal (see [HeSt52] and [Hest80]).

Notice that unlike other methods, we don't need any information about the Hessian matrix of f at x_k when using the conjugate gradient method. Thus the conjugate gradient method may be the only method which is applicable to very large problems, i.e. problems with hundreds or thousands of variables.

The conjugate direction method was first devised for quadratic objective functions by Hestenes and Stiefel [HeSt52] and extended to nonlinear functions by Fletcher and Reeves [FlRe64]. The method was also used in solving large linear system of equations (see [CGOl76], [Axel74]). Nazareth [Naza79] has shown that the conjugate gradient method and **BFGS** method are theoretically identical when applied to quadratic functions with exact line searches. Many variations have been proposed to improve the original conjugate gradient method, e.g. restarting techniques and preconditioned methods to solve systems of linear equations (see [CGOl76], [Axel74] and [WoJi89]). [Hest80] gives a very good survey of the conjugate gradient method.

As we know, Newton's method has locally Q-quadratic rate of convergence but needs to calculate the Hessian of the objective function at each iterative point, which is often rather expensive. Alternatives are new strategies to approximate

the Hessian and to improve the local convergence property to a global convergence property and yet to keep the rapid rate of convergence. Such methods are *quasi-Newton* methods.

ALGORITHM 2.1.4. *(Quasi-Newton method) Given initial point* $x_0 \in \mathbb{R}^n$, *for* $k = 0, 1, 2, \cdots$, *until "convergence",*

1. *calculate* $\nabla f(x_k)$;
2. *compute* B_k *as an approximate to the Hessian of* f *at* x_k;
3. *solve* $B_k s_k = -\nabla f(x_k)$;
4. *decide whether to take the quasi-Newton step* $x_{k+1} := x_k + s_k$ *or choose* x_{k+1} *by global strategy.*

The first quasi-Newton method, called the *variable metric* method, was suggested in [Davi59]. Assume that we have computed $x_{k+1} := x_k + s_k$, $y_k := \nabla f(x_{k+1}) - \nabla f(x_k)$. Then the new Hessian approximate B_{k+1} is obtained by updating B_k by taking account of the newly-acquired information:

$$B_{k+1} s_k = y_k, \tag{2.1.1}$$

which is called the *quasi-Newton equation*. Together with other additional requirements, e.g. the symmetric property of B_{k+1}, we obtain a number of quasi-Newton updating formulae:

$$\begin{aligned} \textbf{PSB:} \qquad B_{k+1} := B_k &+ \frac{1}{s_k^T s_k}\left((y_k - B_k s_k) s_k^T + s_k (y_k - B_k s_k)^T\right) \\ &- \frac{(y_k - B_k s_k)^T s_k}{(s_k^T s_k)^2} s_k^T s_k, \end{aligned} \tag{2.1.2}$$

which is termed the Powell–symmetric–Broyden update, see [Powe70], [DeMo77];

$$\textbf{DFP:} \qquad B_{k+1} := B_k + \frac{(y_k - B_k s_k) y_k^T + y_k (y_k - B_k s_k)^T}{y_k^T s_k} - \frac{(y_k - B_k s_k)^T s_k}{(y_k^T s_k)^2} y_k^T y_k \tag{2.1.3}$$
$$= \left(I - \frac{y_k s_k^T}{y_k^T s_k} \right) B_k \left(I - \frac{s_k y_k^T}{y_k^T s_k} \right) + \frac{y_k y_k^T}{y_k^T s_k},$$

which is termed the Davidon–Fletcher–Powell, see [Davi59], [FlPo63];

$$\textbf{BFGS:} \qquad B_{k+1} := B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k},$$

which is termed the Broyden–Fletcher–Goldfarb–Shanno, see [Broy70], [Flet70], [Gold70] and [Shan70].

Under reasonable conditions on the objective function, one can prove Q-superlinear rate of convergence for quasi-Newton methods using the **DFP** or **BFGS** update, see [BDMo73], [Powe71], [Powe75], [DeMo74] and [Stoe75].

2.2 Line Searches and the Trust Region Approach

Newton's method converges locally. Therefore it is necessary to consider a strategy to make it globally convergent; one example of such a strategy is a line search. The idea of line searches is simple: given a descent direction p_k, we take a step in that direction which yields an *acceptable* x_{k+1}.

ALGORITHM 2.2.1. *(Exact line searches) Given the current iterative point* x_k *and a descent direction* p_k,

1. find λ_k *such that*

$$f(x_k + \lambda_k p_k) = \min\{ f(x_k + \lambda p_k) : \lambda \geq 0 \}; \tag{2.2.1}$$

2. $x_{k+1} := x_k + \lambda_k p_k$.

Until the mid 1960s the prevailing belief was that λ_k should be chosen to solve the one-dimensional minimization problem (2.2.1) accurately. More careful computational tests have led to a complete turnabout, i.e. we can give a weak acceptance criteria for λ_k leading to methods that perform as well in theory and better in practice. The first criterion is that the average rate of decrease from $f(x_k)$ to $f(x_{k+1})$ be at least some prescribed fraction of the initial rate of decrease in that direction, i.e. we pick up an $\alpha \in (0,1)$ and choose $\lambda_k > 0$ such that

$$f(x_k + \lambda p_k) \leq f(x_k) + \alpha\lambda \nabla f(x_k)^T p_k. \tag{2.2.2}$$

The second criterion is that the rate of decrease of f in the direction p_k at x_{k+1} be larger than some prescribed fraction of the rate of decrease in the direction p_k at x_k . That is

$$\nabla f(x_{k+1})^T p_k = \nabla f(x_k + \lambda_k p_k)^T p_k \geq \beta \nabla f(x_k)^T p_k, \tag{2.2.3}$$

for some $\beta \in (\alpha, 1)$, see [Armi66] and [Gold67].

It is proven that the above two criteria are always consistent and the corresponding line search algorithm has the same rate of convergence as the original method (i.e. one of the methods described in §2.1) under reasonable conditions, see [Wolf67] and [Wolf71].

Suppose that we take the descent direction as a quasi-Newton direction. If the full quasi-Newton direction is not satisfactory, then it indicates that our quadratic model doesn't adequately model f in a spherical neighborhood of the current iterative point containing the full quasi-Newton step. In line search algorithms we retain the same step direction and choose a shorter step length. This new length is

based on a one-dimensional quadratic model along the quasi-Newton direction. It has the disadvantage that it makes no further use of the n-dimensional quadratic model. This motivated people to use the trust region approach instead of using line searches, i.e. consider the solution to

$$\min\{\, \phi_k(s) : |D_k s| \le \Delta_k \,\} \tag{2.2.4}$$

as our new step length, where

$$\phi_k(s) := f_k + \nabla f(x_k)^T s + \frac{1}{2} s^T B_k s \tag{2.2.5}$$

is the model function approximating the objective function at x_k, and Δ_k is the trust region radius, and D_k is the scaling matrix which will be discussed in §4.2.

The idea of the model trust region approach was suggested for nonlinear least-square problems by Levenberg [Leve44] and Marquardt [Marq63] and generalized to solving nonlinear optimization problems by Goldfeld, Quandt and Trotter [GQTr66].

ALGORITHM 2.2.2. *(The trust region approach) Given the current iterative point x_k, its corresponding model (2.2.5), and its trust region radius Δ_k, repeat*

1. find s_k, the approximate solution to (2.2.4), calculate $x_{k+1} := x_k + s_k$;

2. decide whether x_{k+1} is acceptable, and calculate a new value of Δ_k

until x_{k+1} is an acceptable next point, then take $\Delta_{k+1} = \Delta_k$.

There are two classes of ways to compute the solution to (2.2.4). The first class is typically like the *dogleg* or *double dogleg* methods suggested in [Powe70] and [DeMe79], which is efficient especially for lower dimensional problems. The second class is based on the methods suggested in [Hebd73] and [More77], and theoretically supported by the following theorem proved in [Sore80] and [Gay81].

THEOREM 2.2.3. *A vector* s^* *is a global solution to problem (2.2.4) iff* $\|D_k s^*\| \leq \Delta_k$, *and for some* $\lambda^* \geq 0$,

$$(B_k + \lambda^* D_k^T D_k)s_* = -g_k, \quad \lambda^*(\|D_k s^*\| - \Delta_k) = 0,$$

with $B_k + \lambda^* D_k^T D_k$ *positive semidefinite. If* $B_k + \lambda^* D_k^T D_k$ *is positive definite, then (2.2.4) has a unique global solution.*

The updating of the trust region radius is very important to the implementation of the trust region algorithm. For details of the various schemes for updating the trust region radius; see [Flet71], [More77], [Gay79], [Flet80] and [Sore80b].

CHAPTER 3

MOTIVATIONS OF THE TRUST REGION APPROACH FOR CONIC MODELS

Many optimization algorithms update the quadratic approximation to the objective function $f(x)$. In the steepest descent direction algorithm, each quadratic approximation has a unit Hessian and matches the gradient of f at the iterative point. In Newton's method, each approximation matches the Hessian as well as the gradient of f at the iterative point. In quasi-Newton algorithms, each approximation matches the gradient of f at two points (see [Davi59], [Huan70] and [DeMo77]), i.e. two latest iterative points, which is equivalent to the quadratic approximation matching the gradient and having an approximate Hessian. It is interesting to generalize the quadratic approximation (model) to other kinds of approximations (models) because the model approximation plays a very important role in unconstrained optimization techniques and is essential for obtaining better model functions.

A good model function should have the following properties: 1) It is easy to formulate the model approximation for each new iterative point. For the quadratic model, we need function, gradient and Hessian (or its approximation) evaluations at the iterative point. These are vital elements in order to approximate the objective function locally; 2) The minimum of the model function should be easy to obtain. In the case of the quadratic model, we need only solve a linear system to obtain the minimum of the model function; 3) The model function has enough degrees of freedom to approximate the objective function, which means more vital information such as the information inherited from the past few iterations should be used in the approximation.

3.1 Conic Models

Davidon [Davi80] proposed conic functions as a candidate of model functions to approximate the objective function. A typical explicit conic function is

$$\psi(s) = f + \frac{g^T s}{1 - a^T s} + \frac{1}{2}\frac{s^T A s}{(1 - a^T s)^2}, \tag{3.1.1}$$

where $a \in \mathbb{R}^n$ such that $1 - a^T s > 0$ for a sufficiently large conical domain in s.

Reasons for proposing such a function as model approximate function are (see [Davi80]):

1) "Under some reasonable conditions, a conic interpolation is determined by successive function and gradient evaluations at $n+1$ vertices of an n dimensional simplex, using $O(n^2)$ numerical operations after each, much as a quadratic interpolation can be determined, to within an additive constant, by just its gradients at these vertices."

2) "Optimization algorithms using conic approximations can be made invariant under the group of collinear transformations characteristic of projective geometry. Newton–Raphson and quasi-Newton algorithms are invariant only under a proper subgroup of affine transformation, while steepest descent and conjugate gradient algorithms are invariant only under the still smaller subgroup of isometries of Euclidean space. While affine transformations can improve the conditioning of the Hessian of the objective function at any point, collinear transformations can also make the transformed Hessian more nearly constant, since the Jacobian of collinear transformations, unlike that of affine ones, need not be constant."

3) "Unlike other nonquadratic approximations considered in [Frie71], [JaOx72] and [Char73], from which the minimizer is always in the direction of a Newton

step," the minimizer of a conic model need not be in the direction of a Newton step.

In order to take a deeper look at the conic functions, we need the following definitions (see [Davi80]).

DEFINITION 3.1.1. *A smooth function* $f : \mathbb{R}^n \to \mathbb{R}$ *is*

affine iff the gradient of f *is a constant;*

quadratic iff its Hessian is a constant;

collinear iff it is a ratio of two affine functions;

conic iff it is a ratio of a quadratic and the square of an affine function.

A map $\phi : \mathbb{R}^m \to \mathbb{R}^n$ *is affine, quadratic, collinear or conic iff any affine function* $f : \mathbb{R}^n \to \mathbb{R}$ *makes the composition* $f \circ \phi : \mathbb{R}^m \to \mathbb{R}$ *affine, quadratic collinear or conic.*

The set of critical points of each conic function $f : \mathbb{R}^n \to \mathbb{R}$ is an affine subspace in $\mathbb{R}^n$, which is possibly null. This point can be seen in the proof of *Lemma 3.1.2*. The level sets of each conic function are conic sections. These conic sections are similar and concentric iff the function is quadratic. To within an affine transformation of its domain, each conic function with a unique minimizer is equal to one with unit Hessian at its minimizer. For each such normalized conic function, there is just one $a \in \mathbb{R}^n$ with

$$a^T x < 1 \quad \text{and} \quad f(x) = f_* + \frac{1}{2}\frac{x^T x}{(1 - a^T x)^2} \tag{3.1.2}$$

for all $x \in \mathbb{R}^n$ and a constant f_*. The level sets of $f(x)$ are conic sections which are invariant under rotations about a and have one *focus* at $0 \in \mathbb{R}^n$. This function is quadratic and its level sets are concentric spheres iff $a = 0$. For

$a \neq 0$, there is just one level set for each eccentricity $e \geq 0$. That is

$$\{ x \in \mathbb{R}^n : a^T a x^T x \leq e^2 (1 - a^T x)^2 \} \tag{3.1.3}$$

within which $f(x) \leq f_* + \frac{1}{2} e^2 / a^T a$. This is an ellipsoid for $e < 1$, a paraboloid for $e = 1$, and a lobe of hyperboloid for $e > 1$. And the function is convex only in the segment of a paraboloid:

$$\{ x \in \mathbb{R}^n : a^T a x^T x \leq (1 + a^T x)^2 \}.$$

From the calculus point of view, we have the following lemma about the representation of conic functions and the necessary and sufficient conditions for the minimum of a conic function. The lemma is from [Davi80], and a detailed proof is provided here because of the importance of the lemma in the study of conic functions. (Note that a proof was not given in [Davi80]).

LEMMA 3.1.2. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is conic and has the value $f_0 \in \mathbb{R}$ and gradient $g_0 \in \mathbb{R}^n$ at x_0 iff there is an $a \in \mathbb{R}^n$ and an $n \times n$ matrix A with*

$$\psi(s) := f(x_0 + s) = f_0 + \frac{g_0^T s}{1 - a^T s} + \frac{1}{2} \frac{s^T A s}{(1 - a^T s)^2}, \tag{3.1.4}$$

for all s such that $a^T s < 1$. The function is collinear iff $A = 0$. The gradient of f at $x = x_0 + s$ or ψ at s can be expressed by

$$\nabla \psi(s) = \frac{1}{\nu^2} (I - a s^T)^{-1} (\nu g_0 + A s) = \frac{1}{\nu^3} (\nu I + a s^T)(\nu g_0 + A s) \tag{3.1.5}$$

where $\nu = 1 - a^T s$. This gradient vanishes at s_* iff

$$(A - g_0 a^T) s_* = -g_0. \tag{3.1.6}$$

The value of $\psi(s)$ is the same at all critical point s_* and equals

$$\psi(s_*) = f_0 + \frac{1}{2} g_0^T s_* / \nu_* = f_0 - \frac{1}{2} s_*^T A s_* / \nu_*^2, \tag{3.1.7}$$

where $\nu_* = 1 - a^T s_*$. *The Hessian of* $\psi(s)$ *at* s_* *is*

$$\begin{aligned} &\frac{1}{\nu_*^2}(A - ag_0^T - g_0 a^T + 2(f_0 - \psi(s_*))aa^T) \\ =&\frac{1}{\nu_*^2}(I - as_*^T)^{-1} A (I - s_* a^T)^{-1} \\ =&\frac{1}{\nu_*^4}(\nu_* I + as_*^T) A (\nu_* I + s_* a^T). \end{aligned} \tag{3.1.8}$$

A critical point is a minimizer iff $A \geq 0$ *(positive semidefinite), and it is a unique minimizer iff* $A > 0$ *(positive definite), in which case,* $\nu_* = 1/(1 - a^T A^{-1} g_0)$, $s_* = -\nu_* A^{-1} g_0$, $\psi(s_*) = f_0 - \frac{1}{2} g_0^T A^{-1} g_0$, *and the Hessian of* ψ *at* s_* *is*

$$\frac{1}{\nu_*^2}(A - ag_0^T) A^{-1} (A - g_0 a^T). \tag{3.1.9}$$

PROOF: One can easily check (3.1.4)–(3.1.8). It is important to emphasize that the value of $\psi(s)$ is the same at all critical points s_* because the convex combination of critical points is again a critical point.

Now suppose $A \geq 0$. Given $h \in \mathbb{R}^n$, we consider $\psi(s_* + h)$. As $(I - s_* a^T)$ is invertible, we can denote $w := (I - s_* a^T)^{-1} h$. Then $w \neq 0$ iff $h \neq 0$, and

$$h = (I - s_* a^T) w = w - a^T w s_*$$

$$s_* + h = (1 - a^T w) s_* + w$$

$$1 - a^T (s_* + h) = (1 - a^T w)(1 - a^T s_*)$$

$$g_0^T (s_* + h) = (1 - a^T w) g_0^T s_* + g_0^T w.$$

So we have

$$\begin{aligned}
\psi(s_* + h) &= f_0 + \frac{(1 - a^T w) g_0^T s_* + g_0^T w}{(1 - a^T w)(1 - a^T s_*)} \\
&\quad + \frac{1}{2} \frac{(1 - a^T w)^2 s_*^T A s_* + 2(1 - a^T w) s_*^T A w + w^T A w}{(1 - a^T w)^2 (1 - a^T s_*)^2} \\
&= f_0 + \frac{g_0^T s_*}{(1 - a^T s_*)} + \frac{1}{2} \frac{s_*^T A s_*}{(1 - a^T s_*)^2} \\
&\quad + \frac{g_0^T w (1 - a^T w)(1 - a^T s_*) + (1 - a^T w) s_*^T A w + \frac{1}{2} w^T A w}{(1 - a^T w)^2 (1 - a^T s_*)^2} \\
&= f_0 + \frac{g_0^T s_*}{(1 - a^T s_*)} + \frac{1}{2} \frac{s_*^T A s_*}{(1 - a^T s_*)^2} \\
&\quad + \frac{(1 - a^T w)\big(s_*^T A + (1 - a^T s_*) g_0^T\big) w + \frac{1}{2} w^T A w}{(1 - a^T w)^2 (1 - a^T s_*)^2} \\
&= f_0 + \frac{g_0^T s_*}{(1 - a^T s_*)} + \frac{1}{2} \frac{s_*^T A s_*}{(1 - a^T s_*)^2} \\
&\quad + \frac{1}{2} \frac{w^T A w}{(1 - a^T w)^2 (1 - a^T s_*)^2} \\
&= \psi(s_*) + \frac{1}{2} \frac{w^T A w}{(1 - a^T w)^2 (1 - a^T s_*)^2} \\
&\geq \psi(s_*),
\end{aligned}$$

i.e. s_* is a minimizer iff $A \geq 0$, and a unique minimizer iff $A > 0$. One can easily check (3.1.9). □

3.2 Collinear Scaling Conic Models

If we take the model approximation to the objective function as interpolation, then the quadratic model doesn't possess enough degrees of freedom to meet the interpolation conditions we desire. However we don't wish to abandon the quadratic model because a large amount of practical experience and theoretical work indicate its suitability. In order to increase the degrees of freedom, yet retain

the desirable properties of the quadratic model, collinear scaling

$$s = \frac{Jw}{1 + h^T w} \tag{3.2.1}$$

was proposed in [Sore80] to scale the objective function $\phi(w) := f(x(w))$, where $J \in \mathbb{R}^{n \times m}$, $h \in \mathbb{R}^m$, $w \in \mathbb{R}^m$, $x(w) := x + s = x + Jw/(1 + h^T w)$. A method based on a local quadratic model $\psi(w)$ to the scaled function $\phi(w)$ was developed. The quadratic approximation takes the following form:

$$\psi(w) = \phi(0) + \nabla\phi(0)^T w + \frac{1}{2} w^T B w. \tag{3.2.2}$$

By using the chain rule, we have

$$\psi(w) = f(x) + \nabla f(x)^T J w + \frac{1}{2} w^T B w. \tag{3.2.3}$$

If B is symmetric and positive definite, then the step v that solves

$$Bv = -J^T \nabla f(x)$$

is a predicted minimizer of the scaled function $\phi(w)$. The step s from x to $\bar{x}$ is

$$s = \frac{Jv}{h^T v + 1},$$

i.e. $\bar{x} := x(v)$, which is a new estimate of the solution. As before, if $\bar{x}$ is an acceptable approximation to a local minimizer of f then we wish to update the collinear scaling model and also the quadratic model of the new scaled function $\bar{\phi}(w) = f(\bar{x}(w))$ with

$$\bar{x}(w) = \bar{x} + \frac{\bar{J} w}{\bar{h}^T w + 1}. \tag{3.2.4}$$

For the purpose of interpolating the objective function with function values and first order derivatives at both x and $\bar{x}$, the following interpolating conditions

are developed in [Sore80]:

$$f(x) = f(\bar{x}) - \gamma \nabla f(\bar{x}) s + \frac{1}{2} v^T B v, \tag{3.2.5}$$

$$\frac{\nabla f(x)}{\gamma}(\bar{J} + s\bar{h}^T) = \nabla f(\bar{x})\bar{J} - v^T \bar{B},$$

where $\gamma = -g^T s/(\rho + f - \bar{f})$ with $\rho = \left((f - \bar{f})^2 - (g^T s)(\bar{g}^T s)\right)^{\frac{1}{2}}$. Denoting

$$r = \bar{J}^T g - \frac{1}{\gamma}(\bar{J} + s\bar{h}^T)^T g,$$

we have the *generalized quasi-Newton equations*:

$$\bar{B}v = r; \qquad \bar{J}v = \gamma s; \qquad \bar{h}^T v = 1 - \gamma, \tag{3.2.6}$$

which are central to the theory of quasi-Newton methods.

The updating of the model includes the updating of J to $\bar{J}$, h to $\bar{h}$ and B to $\bar{B}$. It is desirable that the past information during recent iterations be preserved. For example, if W is the linear span of previous scaled search directions and let $W := span\{W, v\}$, then a natural requirement is that

$$\bar{\phi}(w - v) = \phi(w) \tag{3.2.7}$$

for all $w \in W' \subset W$, with $W' := \{\, w \in W : h^T w + 1 > 0 \,\}$, an open convex neighborhood of $0 \in W$. Condition (3.2.7) leads immediately to the requirement

$$\bar{x}(w - v) = x + \frac{(\bar{J} + s\bar{h}^T)w}{\bar{h}^T w + \gamma}, \tag{3.2.8}$$

for all $w \in W'$, which, together with the quasi-Newton formula (3.2.6), leads to the following updating formula for J and h (see [Sore80]):

$$\bar{J} := \gamma(J - sh^T), \tag{3.2.9}$$

$$\bar{h} := \gamma Q h + \Big(\frac{1 - \gamma - \gamma v^T Q h}{v^T P d}\Big) P d \tag{3.2.10}$$

for any vector d such that $v^T P d \neq 0$, where Q is an orthogonal projection onto W and $P = I - Q$.

For the updating of B, [Schn77] proposed general requirements:

$$\bar{B} \quad \text{symmetric}, \qquad \bar{B}v = r, \qquad (\bar{B} - B)w = 0 \quad \text{for all} \quad w \in W. \tag{3.2.11}$$

Therefore, it is natural to use the **DFP** or **BFGS** updating formula. In this way, a conic model algorithm with line searches was formulated.

ALGORITHM 3.2.1. *([Sore80], Conic model algorithm with line searches)*

Step 0. *Set* $k := 1$, *initialize* x_k, J_k, h_k *and* B_k *positive definite. Put* $Q_k := 0$, $P_k := I$, $Q'_k := 0$, *evaluate* $g_k := \nabla f(x_k)$, $f_k := f(x_k)$;

Step 1. *Compute* $v_k := -B_k^{-1} J_k^T g_k$ *and search* $f(x_k(\alpha))$, *where*

$$x_k(\alpha) := x_k + \frac{J_k(\alpha v_k)}{h_k^T(\alpha v_k) + 1},$$

until an $\alpha > 0$ *has been found such that*

$$\rho := (f_k - f_{k+1})^2 - (g_k^T s_k)(g_{k+1}^T s_k) \geq 0,$$

with $x_{k+1} := x_k(\alpha)$, $s_k := x_{k+1} - x_k$, $f_{k+1} := f(x_{k+1})$, $g_{k+1} := \nabla f(x_{k+1})$. *Compute* $\rho := \rho^{\frac{1}{2}}$ *and*

$$\gamma := -g_k^T s_k / (\rho + f_k - f_{k+1}),$$

and put $v := \alpha v$;

Step 2. *If convergence occurs, then stop. Otherwise*

(a) $J_{k+1} := \gamma(J_k - s_k h_k^T)$;

(b) Choose a vector d_k *and put*

$$h_{k+1} := \gamma Q_k h_k + \Big(\frac{1 - \gamma - \gamma v_k^T Q_k h_k}{v_k^T P_k d_k}\Big) P_k d_k; \tag{3.2.12}$$

(c) $r := J_{k+1}^T g_{k+1} - (1/\gamma)(J_{k+1} + s_k h_{k+1}^T)^T g_k$;

(d) Updating B_k *to* B_{k+1} *such that (3.2.11) are satisfied;*

(e) $q := P_k v_k / |P_k v_k|$, $Q_{k+1} := (Q_k, q)$, $Q'_{k+1} := Q_{k+1} Q_{k+1}^T$, $P_{k+1} := I - Q_{k+1}$;

(f) go to **step 1**.

Under reasonable conditions on the objective function and some desirable restrictions on the implementation, it's proved that the class of algorithms has Q-linear rate of convergence. Under further conditions, it is proved that the class of algorithms has Q-superlinear rate convergence, which is the best result that this kind of algorithm can achieve for solving unconstrained minimization problems without using analytic Hessian. A specific algorithm using a conic model with line searches was developed, and the computational results show that the algorithm has very good performance on some of the standard test functions; see [Sore80].

3.3 The Trust Region Approach for Conic Models and its Subproblems

Conic model methods proposed by Sorensen use line searches. Therefore once a model is formulated, the minimization is limited to the descent direction determined by the model, i.e. a one-dimensional minimization along a fixed descent direction at each iteration, which is an inadequate use of the model as the model is always a good approximation in a certain n-dimensional neighborhood of the iterative point. Therefore the trust region approach is an alternative to implement conic model algorithms. All the reasons why we use the trust region approach for quadratic models apply to conic models; see [DeMe79], [Flet80], [More82], [Sore82] and [DeSc83].

The trust region approach for quadratic models is well-developed, and the corresponding numerical algorithms have already been adopted in standard mathematical software packages (See [Gay83] and [IMSL10]). To develop the trust region approach for conic models, it's important to establish the corresponding result to *Theorem 2.2.3* for conic models, i.e. to establish the analytic solution to the trust region subproblems of conic models.

Suppose $f(x)$ is the objective function. We use a model function $\psi(s)$ to approximate $f(x_k + s)$ at the k-th iteration. We are interested in two kinds of conic models. The first kind is a class of explicit conic functions:

$$\psi(s) = f + \frac{g^T s}{(1 - a^T s)} + \frac{1}{2}\,\frac{s^T A s}{(1 - a^T s)^2}, \tag{3.3.1}$$

proposed in [Davi80], where $f := f(x_k)$, $g := \nabla f(x_k)$, a and A are other model updating informations. The second kind is a class of collinear scaling conic functions:

$$\begin{cases} \psi(s) &= f + g^T J w + \frac{1}{2} w^T B w, \\ s &= Jw/(1 + h^T w), \end{cases} \tag{3.3.2}$$

proposed in [Sore80], where h, J, B are other model updating informations. Their trust region subproblem is

$$\min\{\psi(s) \mid \|Ds\| \leq \Delta\}, \tag{3.3.3}$$

where D is the scaling matrix, Δ is the trust region radius.

We present a necessary and sufficient condition for the solution to problem (3.3.3) for both kinds of conic functions. This solution plays an important role in the trust region method for the conic models, both theoretically and numerically.

THEOREM 3.3.1. *For conic model functions (3.3.1), suppose* $g \neq 0$, $A \geq 0$, *and* $1 - a^T s > 0$ *for all* $s \in \mathbb{R}^n$ *satisfying* $\|Ds\| \leq \Delta$. *If* s_* *solves (3.3.3)*

and $D^T DS_* \neq 0$, *then there exists* $\mu^* \geq 0$, *such that*

$$(A - ga^T + \mu^* D^T D)s_* = -g + \mu^* \Delta^2 a, \tag{3.3.4}$$

$$\mu^*(\|Ds_*\| - \Delta) = 0 \tag{3.3.5}$$

PROOF: Let $\eta(s) = \|Ds\| - \Delta$, then the gradient of $\eta(s)$ is

$$\nabla\eta(s) = D^T Ds/\|Ds\|, \quad \text{(provide} \quad \|Ds\| \neq 0).$$

From *Lemma 3.1.2*,

$$\nabla\psi(s) = \frac{1}{\nu^2}(I - as^T)^{-1}(\nu g + As),$$

where $\nu = 1 - a^T s$. By the Lagrange multiplier rule and the Kuhn-Tucker constraint qualification (see [Avri76], Chapter 3), if s_* solves (3.3.3) then there exists $\mu \geq 0$ such that

(i) $\nabla\psi(s_*) + \mu\nabla\eta(s_*) = 0$,

(ii) $\mu\eta(s_*) = 0$.

Then

$$\begin{aligned} &\nu_* g + As_* + \mu\nu_*^2(I - as_*^T)D^T Ds_*/\|Ds_*\| = 0, \\ &\mu(\|Ds_*\| - \Delta) = 0. \end{aligned} \tag{3.3.6}$$

Note that (3.3.5) follows immediately from (3.3.6). By assumption, $g \neq 0$ and $1 - a^T s > 0$ for all $s \in \mathbb{R}^n$ satisfying $\|Ds\| \leq \Delta$, so we have $\|s_*\| \neq 0$, and $\nu_* = 1 - a^T s_* > 0$. Denote $\mu^* = \mu\nu_*^2/\|Ds_*\|$, then $\mu^* \geq 0$, and

$$(1 - a^T s_*)g + As_* + \mu^* D^T Ds_* - \mu^*\|Ds_*\|^2 a = 0. \tag{3.3.7}$$

By (3.3.6), if $\|Ds_*\| < \Delta$, then $\mu^* = 0$, and (3.3.4) is obvious from (3.3.7). Otherwise $\|Ds_*\| = \Delta$, and (3.3.7) becomes

$$g - ga^T s_* + As_* + \mu^* D^T Ds_* - \mu^*\Delta^2 a = 0, \tag{3.3.8}$$

and (3.3.4) again follows. □

THEOREM 3.3.2. *For conic model functions (3.3.1), let* $g \neq 0$, $A \in \mathbb{R}^{n\times n}$, $A \geq 0$ *and* $1 - a^T s > 0$ *for all* $s \in \mathbb{R}^n$ *such that* $\|Ds\| \leq \Delta$. *If there exist* $\mu \geq 0$, $p \in \mathbb{R}^n$ *such that* $\|Dp\| \leq \Delta$, *and*

$$(A - ga^T + \mu D^T D)p = -g + \mu\Delta^2 a, \tag{3.3.9}$$

$$\mu(\|Dp\| - \Delta) = 0, \tag{3.3.10}$$

with $\bar{A} := A + \mu D^T D - \mu\Delta^2 aa^T \geq 0$, *then* p *solves (3.3.3). The solution is unique iff* $\bar{A} > 0$.

PROOF: If $\mu = 0$, then p solves

$$(A - ga^T)p = -g. \tag{3.3.11}$$

Since $g \neq 0$, $A \geq 0$, and $\|Dp\| \leq \Delta$, p solves (3.3.3) by *Lemma 3.1.2*.

Now suppose $\mu > 0$. Define

$$\psi_\mu(s) \equiv f + \frac{(g - \mu\Delta^2 a)^T s}{(1 - a^T s)} + \frac{1}{2}\,\frac{s^T(A + \mu D^T D - \mu\Delta^2 aa^T)s}{(1 - a^T s)^2}. \tag{3.3.12}$$

Because $A+\mu D^T D-\mu\Delta^2 aa^T \geq 0$, it follows from *Lemma 3.1.2* that the solution of the equation

$$(A + \mu D^T D - \mu\Delta^2 aa^T - (g - \mu\Delta^2 a)a^T)s = -(g - \mu\Delta^2 a), \tag{3.3.13}$$

i.e.

$$(A + \mu D^T D - ga^T)s = -(g - \mu\Delta^2 a),$$

is a global minimizer of $\psi_\mu(s)$, so p is a global minimizer of $\psi_\mu(s)$. The point p is a unique minimizer iff $\bar{A} = (A + \mu D^T D - \mu\Delta^2 aa^T) > 0$. So we have

$$\psi_\mu(s) \geq \psi_\mu(p), \quad \text{for all} \quad s \quad \text{s.t.} \ 1 - a^T s > 0, \tag{3.3.14}$$

and the inequality is strict for all $s \neq p$ iff $\bar{A} > 0$. Now,

$$\psi(s) = \psi_\mu(s) + \overline{\psi}(s),$$

where

$$\overline{\psi}(s) = \frac{\mu\Delta^2 a^T s}{1 - a^T s} - \frac{1}{2}\,\frac{\mu s^T D^T D s}{(1 - a^T s)^2} + \frac{1}{2}\,\frac{\mu\Delta^2 s^T a a^T s}{(1 - a^T s)^2}.$$

So we have

$$\psi(s) \geq \psi(p) + \overline{\psi}(s) - \overline{\psi}(p), \quad \text{for all } s \quad \text{s.t. } 1 - a^T s > 0. \tag{3.3.15}$$

Consider $s \in \mathbb{R}^n$ with $\|Ds\| \leq \Delta$. By assumption $1 - a^T s > 0$ for any such s, therefore,

$$\begin{aligned}
\overline{\psi}(s) &\geq \frac{\mu\Delta^2 a^T s}{1 - a^T s} - \frac{1}{2}\,\frac{\mu\Delta^2}{(1 - a^T s)^2} + \frac{1}{2}\,\frac{\mu\Delta^2 (a^T s)^2}{(1 - a^T s)^2} \\
&= -\frac{1}{2}\frac{\mu\Delta^2}{(1 - a^T s)^2}(-2a^T s(1 - a^T s) + 1 - (a^T s)^2) \\
&= -\frac{1}{2}\,\frac{\mu\Delta^2}{(1 - a^T s)^2}(1 - 2a^T s + (a^T s)^2) \\
&= -\frac{1}{2}\mu\Delta^2.
\end{aligned} \tag{3.3.16}$$

Since by (3.3.10) $\|Dp\| = \Delta$ if $\mu \neq 0$, by the same analysis as above,

$$\overline{\psi}(p) = -\frac{1}{2}\mu\Delta^2. \tag{3.3.17}$$

So from (3.3.15), (3.3.16) and (3.3.17),

$$\begin{aligned}
\psi(s) &\geq \psi(p) + \overline{\psi}(s) - \overline{\psi}(p) \\
&\geq \psi(p) + \frac{(-1)}{2}\mu\Delta^2 - \frac{(-1)}{2}\mu\Delta^2 \\
&= \psi(p) \qquad \text{for all } s \in \mathbb{R}^n \quad \text{s.t. } \|Ds\| \leq \Delta.
\end{aligned} \tag{3.3.18}$$

Again from *Lemma 3.1.2*, the inequality (3.3.18) is strict for $s \neq p$ iff $\bar{A} > 0$, i.e. p is the unique solution iff $A + \mu D^T D - \mu\Delta^2 aa^T$ is positive definite. □

Theorem 3.3.1 and *Theorem 3.3.2* provide necessary and sufficient optimality conditions for the solution to the subproblem (3.3.3) for conic function (3.3.1). For the collinear scaling conic model (3.3.2), we can prove analogous results by using the same techniques as in *Theorem 3.3.1* and *Theorem 3.3.2*.

THEOREM 3.3.3. *For conic model functions (3.3.2), let J^+ be the Moore–Penrose generalized inverse of J. Denote $\bar{B} = {J^+}^T B J^+$, and*

$$\tilde{B} = \bar{B} - \mu D^T D + \mu\Delta^2 {J^+}^T h h^T J^+. \tag{3.3.19}$$

Assume $(JJ^+)^T g \neq 0$, $1 - h^T J^+ s > 0$. for each $s \in \mathbb{R}^n$, with $\|Ds\| \leq \Delta$, and $\bar{B} \geq 0$. If there exist $\mu \geq 0$ and $p \in \mathbb{R}^n$ such that $\|Dp\| \leq \Delta$, $\tilde{B} \geq 0$, and

$$(\bar{B} - JJ^+ g h^T J^+ + \mu D^T D)p = -JJ^+ g + \mu\Delta^2 {J^+}^T h, \tag{3.3.20}$$

$$\mu(\|Dp\| - \Delta) = 0 \tag{3.3.21}$$

then p solves (3.3.3) for $\psi(s)$ defined in (3.3.2), and the solution is unique iff $\tilde{B}$ is positive definite. If p solves (3.3.3) for $\psi(s)$ defined in (3.3.2) and $D^T Dp \neq 0$, then there exists $\mu \geq 0$ such that (3.3.20) and (3.3.21) hold.

PROOF: Comparing (3.3.2) with (3.3.1) by setting $a = {J^+}^T h$, $g = (JJ^+)^T g$, $A = {J^+}^T B J^+$, we can prove the theorem by combining *Theorem 3.3.1* and *Theorem 3.3.2*. □

For any specific choices of a in (3.3.1) and h, J in (3.3.2), we can solve the corresponding trust region subproblems numerically by using the above theorems.

We can develop a method similar to the dogleg and double dogleg methods (see [DeMe79]) for each of the situations in *Theorem 3.3.2* and *Theorem 3.3.3*, or we can solve the equations (3.3.9), (3.3.20) by an iterative method using Newton's formula (see §4.4 in [GMWr81]) to get more accurate approximate solutions to the model trust region subproblems.

CHAPTER 4

NUMERICAL IMPLEMENTATION AND COMPUTATIONAL RESULTS

We proposed conic model methods with a trust region approach to solve unconstrained minimization problems in **Chapter 3**. Usually once a new class of numerical methods is considered, good motivations and solid background are established. But the new class of methods can draw attention only after a good numerical performance has been demonstrated and/or some significant theoretical results have been proved. In this chapter, we develop a specific algorithm using a conic model with a trust region approach, display its numerical results and compare the algorithm with other unconstrained minimization solvers.

4.1 Numerical Implementation

We consider a trust region model (3.3.2) which is used to approximate the original function f at the k-th iterative point x_k. Like any typical numerical method for solving unconstrained optimization problems, conic model methods with a trust region approach can be summarized as follows.

ALGORITHM 4.1.1. *(A class of conic model methods with a trust region approach)*

Step 0. *Set* $k := 1$, *initialize* Δ_k, x_k, J_k, h_k, D_k *and* B_k, *calculate* $g_k := \nabla f(x_k)$, $f_k := f(x_k)$;

Step 1. *Solve the following trust region subproblem:*

$$\begin{cases} \min & f_k + g_k^T J_k w + \frac{1}{2} w^T B_k w, \\ & s = J_k w/(1 + h_k^T w), \quad \|D_k s\| \le \Delta_k; \end{cases} \tag{4.1.1}$$

Step 2. *Check if the solution* s_k *to the trust region subproblem is a "successful" one or not. If not, we contract the trust region radius* Δ_k *and go to* **step 1**.

Step 3. *Set* $x_{k+1} := x_k + s_k$, $g_{k+1} := \nabla f(x_{k+1})$, $f_{k+1} := f(x_{k+1})$. *If "convergence" occurs, then stop. Otherwise update* Δ_k , x_k , J_k , h_k , D_k *and* B_k . *Let* $k := k+1$ *and go to* **step 1**.

For the purpose of clarity, we denote the updates of Δ , x , f , g , J , h , B by $\bar{\Delta}$, $\bar{x}$, $\bar{f}$, $\bar{g}$, $\bar{J}$, $\bar{h}$, $\bar{B}$ respectively. Suppose $s = \bar{x} - x$, and v satisfies $s = Jv/(1 + h^T v)$. For J and h , a set of updates which have strong modeling properties due to more degrees of freedom of the conic models, are expressed in (3.2.9) and (3.2.10). For the purpose of simplicity of our discussion, we neglect the projection technique, i.e. we let $P = I$ and $Q = 0$, then (3.2.9) and (3.2.10) are equivalent to

$$\bar{J} = \gamma (J - sh^T); \qquad \bar{h} = \frac{1-\nu}{v^T d}\, d. \tag{4.1.2}$$

The choice of d is such that $v^T d \neq 0$. In [Sore80], d was chosen to be $\bar{J}^T g$. For stability of our algorithm, which will be shown in the proof of the convergence results in **Chapter 5**, we choose $d = \bar{J}^T \bar{g}$.

For B , we use the inverse **BFGS** updating formula to update the inverse Hessian approximation H , i.e. $H = B^{-1}$ and $\bar{H} = \bar{B}^{-1}$. One way of writing this updating formula is

$$\bar{H} = \left(I - \frac{vr^T}{v^T r}\right) H \left(I - \frac{rv^T}{v^T r}\right) + \frac{vv^T}{v^T r}. \tag{4.1.3}$$

Together with the updating formula of J in (4.1.2), we can combine the updating of H and J into a single updating of the matrix JHJ^T .

THEOREM 4.1.2. *By choosing* $d = \bar{J}^T\bar{g}$ *in (4.1.2), we have*

$$\bar{J} = \gamma(H - sh^T); \qquad \bar{h} = \frac{1-\gamma}{\gamma\bar{g}^T s}\bar{J}^T\bar{g}.$$

Then

$$\bar{J}\left(I - \frac{vr^T}{v^T r}\right) = \gamma\left(I - \frac{sy^T}{s^T y}\right)J,$$

where $y = \gamma\bar{g} - \frac{1}{\gamma}g$. *And*

$$\bar{J}\bar{H}\bar{J}^T = \gamma^2\left[\left(I - \frac{sy^T}{s^T y}\right)JHJ^T\left(I - \frac{ys^T}{s^T y}\right)\right] + \frac{ss^T}{s^T y}. \tag{4.1.4}$$

PROOF: Due to the quasi-Newton equation (3.2.6), we have $v^T r = s^T y$, $\bar{J}v = \gamma s$. So

$$(\bar{h}s^T)\bar{g} = (\frac{1-\gamma}{\gamma})\bar{J}^T\bar{g}.$$

It follows that

$$r = \bar{J}^T\bar{g} - \frac{1}{\gamma^2}\bar{J}^T g = (J - sh^T)^T y.$$

Thus

$$\bar{J}\left(I - \frac{vr^T}{v^T r}\right) = \gamma(J - sh^T) - \frac{\gamma s(J^T y - hs^T y)^T}{s^T y} = \gamma\left(I - \frac{sy^T}{s^T y}\right)J,$$

and (4.1.4) follows. □

Denote $C := JHJ^T$, and $\overline{C} := \bar{J}\bar{H}\bar{J}^T$, then (4.1.4) becomes

$$\overline{C} := \gamma^2\left[\left(I - \frac{sy^T}{s^T y}\right)C\left(I - \frac{ys^T}{s^T y}\right)\right] + \frac{ss^T}{s^T y}. \tag{4.1.5}$$

And then we update only C instead of updating both H and J . We keep C in factored form by using Gill and Murray's technique (see [GiMu72], the only modification here is the multiplication of the factor γ^2).

Denote $\alpha := (1-\gamma)/\gamma g^T s$, so that $\bar{h} = \alpha\bar{J}^T\bar{g}$. For convenience of theoretical discussions, we assume J to be an $n \times n$ invertible matrix. Therefore,

if H is invertible, then we have that C is invertible. Denote by $v = Jw$, the subproblem in the k-th iteration becomes

$$\begin{cases} \min & f_k + g_k^T v + \frac{1}{2} v^T C_k^{-1} v, \\ & s = v/(1 + \alpha_k g_k^T v), \quad \|D_k s\| \le \Delta_k; \end{cases} \tag{4.1.6}$$

By *Theorem 3.3.3*, it is equivalent to find μ and s such that

$$(C_k^{-1} - \alpha_k g_k^T g_k + \mu D_k^T D_k)s = -(1 - \mu\alpha_k \Delta_k^2) g_k,$$

$$\mu(\|D_k s\| - \Delta_k) = 0.$$

We use the method similar to *Algorithm (3.4)* in [MoSo83] to get the numerical solution to the subproblem. For the updating of the trust region radius, we refer to the techniques in [More82]. The convergence criteria will be discussed later. A good choice of D_k is rather difficult, but we can choose D_k to be the diagonal matrix which will be discussed in §4.2. Then the algorithm can be specified as follows.

ALGORITHM 4.1.3. *(A specified conic model algorithm with the trust region approach)*

Step 0. *Set* η_1, η_2, ν_1, ν_2, α_0, ζ_1, ζ_2, $\bar{\Delta}$ *such that* $0 < \eta_1 < \frac{1}{2}$, $\frac{1}{2} < \nu_1 < 1 < \nu_2 < 4$, $0 < \alpha_0 < 1$, $0 < \zeta_1 < 1 < \zeta_2 < 2$, *and* $\bar{\Delta} > 0$. *Set* $k := 1$, *initialize* Δ_k, x_k, *and* C_k *positive definite. Calculate* $g := g_k := \nabla f(x_k)$, $f := f_k := f(x_k)$;

Step 1. *Solve the trust region subproblem (4.1.6) with solution* s *and* v. *Let* $x' := x_k + s$, $f' := f(x')$, $g' := \nabla f(x')$. *If* $f' \ge f$, *or* $\rho := (f - f')^2 - (g^T s)(g'^T s) \le 0$, *then let* $\Delta_k = \zeta_1 \Delta_k$, *and go to* **step 1**;

Step 2. *If "convergence" occurs, then stop;*

Step 3. $\rho := \rho^{\frac{1}{2}}$, $\nu := -g^T s/(f - f' + \rho)$, $\nu := \max\{\nu, \nu_1\}$, $\nu := \min\{\nu, \nu_2\}$, $y := \nu g' - \frac{1}{\nu} g$, *update* C_k *to* C_{k+1} *by (4.1.5), and update*

D_k . *Assign*

$$\alpha := sign(\frac{1-\nu}{\nu\, g'^T s}) \min\{\, |\frac{1-\nu}{\nu\, g'^T s}|, \quad \frac{\alpha_0}{\Delta_k \|g_k\|}, \quad \tilde{\alpha}\,\}. \tag{4.1.7}$$

Step 4. *Update the trust region radius:*

a. $pred := -g_k^T v + \frac{1}{2} v^T C_k^{-1} v$, $ared := f - f'$;

b. $PRELDF := ared/pred$;

c. *If* $PRELDF < \eta_1$; *then* $\Delta_k = \zeta_1 \Delta_k$ *and go to* **step 2**;

d. *If* $PRELDF \geq \eta_1$, *then* $\Delta_k := \zeta_2 \Delta_k$;

e. $\Delta_k = \min\{\Delta_k, \bar{\Delta}\}$;

Step 5. *Set* $x_{k+1} := x'$, $f := f_{k+1} := f'$, $g := g_{k+1} := g'$, $k := k+1$ *and go to* **step 1**.

We set bounds for ν , α and Δ_k because of the stability requirement which will be shown in the proofs of the convergence results. The choice of the positive number $\bar{\Delta}$ is pretty free.

4.2 Testing and Computational Results

Testing is essential for computer programs that solve practical minimization problems. Before designing testing, we should first determine what kind of stopping strategy and whether or not scaling should be used.

Stopping is how to determine when to terminate an iterative algorithm and to decide whether the algorithm succeeded (*converged*) or failed. One obvious stopping criterion is to check if

$$|\nabla f(x_k)| \leq \epsilon_1, \tag{4.2.1}$$

because a necessary condition for x^* to solve (2.0.1) is $\nabla f(x^*) = 0$. If (4.2.1) is satisfied, we call the iteration *gradient value convergent*. Although $\nabla f(x) = 0$ can also occur at local maxima or saddle points, globally convergent strategies make convergence virtually impossible to a local maximum or saddle point.

There are many practical problems where we know the optimal value, e.g. least squares problems sometimes have the optimal value 0. In that situation, without loss of generality, another obvious stopping criterion is

$$|f(x_k)| \leq \epsilon_2. \tag{4.2.2}$$

We call this the *absolute value convergence* criterion.

From the numerical point of view, neither of the above criteria is good enough for all kinds of practical problem. For example, (4.2.1) and (4.2.2) don't reflect how close x_k is to x^*, nor do they reflect how close $f(x_k)$ is to $f(x^*)$. In practice, when the iteration is really slow, either in terms of x_k or in terms of $f(x_k)$, we should consider halting the implementation. Therefore the following criteria are more reasonable:

$$PRELDF \leq \epsilon_3, \tag{4.2.3}$$

$$PRELDX \leq \epsilon_4, \tag{4.2.4}$$

where

$$PRELDF := \frac{|f(x_k) - f(x_{k+1})|}{|f(x_k)| + |f(x_{k+1})|}, \tag{4.2.5}$$

$$PRELDX := \frac{|x_k - x_{k+1}|}{|x_k| + |x_{k+1}|}. \tag{4.2.6}$$

As in most iterative procedures, imposing a limit on the number of iterations is always essential. This is largely because we have only limited time, money and

patience. In real applications this limit is determined by the cost of each iteration, which can be high if the function, gradient, and Hessian evaluations are expensive.

In conclusion, each of the above stopping criteria has a different meaning and an individual problem usually has its own stopping requirement. Usually it is a good choice to use a combination of the above stopping criteria.

Scaling is how to adjust for problems that are badly scaled in the sense that the dependent or independent variables are of widely differing magnitudes. For example, we might have a minimization problem in which the first independent variable, x_1 , is in the range of 10^1 to 10^3 meters, and the second independent variable, x_2 , is in the range of 10^{-7} to 10^{-5} seconds. If we use the Euclidean norm $\|x_{k+1} - x_k\|$ of $x_{k+1} - x_k$ in the finite dimensional space to evaluate the improvement of an iteration, it doesn't reflect the improvement of the second variable, x_2 . What we can do about the problem is to rescale the variables, e.g. to choose a diagonal matrix $D_k = diag(10^{-3}, 10^5)$, and consider $\|D_k(x_{k+1} - x_k)\|$ instead of $\|x_{k+1} - x_k\|$ so that each variable of x will have roughly the same magnitude. Although the motivation is really common sense, it is difficult to choose D_k in practice because sometimes we don't know the actual scales of the different variables. Scaling is not limited to the iteration improvement evaluation. It is used throughout every iteration; see [Luen73], [DeSc83] for the details.

Testing is to see whether the computer program works correctly and to see how it compares with other software that solves the same kind of problems. There are many factors considered to be important in justifying the performance of a program, but reliability and cost are considered the most important. By reliability, we mean the ability of the program to successfully solve the problems it is intended for. This is determined by its results on test problems. There is a list of standard test problems for unconstrained minimization which is still be-

ing developed. Cost is usually measured by the running time of the program for test problems. In unconstrained minimization, cost is measured by the number of function, gradient and Hessian (if needed) evaluations the program requires to solve the problems since the **CPU** time is mostly consumed by those evaluations. Sometimes those evaluations should be done by experiments (see §1.2 in [DeSc83]). For large problems, computer storage is also an important issue.

Our algorithm is called **CONTR**. We compare this algorithm with **SUMSL**, (see [Gay83]), which was later adopted by **IMSL** (*International Mathematical and Statistical Library*, the most widely used numerical software), **Version 10**, and is one of the most efficient algorithms for solving unconstrained nonlinear optimization problems. **SUMSL** uses the **BFGS** updating formula to approximate the Hessian with the trust region approach which is carried out by double–dogleg technique in [DeMo79]. **CONTR** and **SUMSL** were tested for all the standard unconstrained nonlinear optimization problems in [MGHi81]. The numerical comparison for $D = I$, which means there is no matrix scaling, is presented in Table 1 and Table 2. All the computations in this section have been carried out on an **Amdahl 5870** computer, using double precision (about 14 decimal places).

The columns in these tables have the following meanings: **LS** is the common logarithm of parameter factor to the standard initial starting point, i.e. we tested each problem with different initial starting point (if it is necessary for the problem), which is 10^{LS} times the original initial starting point; **N** is the problem dimension n ; **NF** is the number of function evaluations performed; **NG** is the number of gradient evaluations; **C** is code of stopping criteria which is explained below; **F** is the final objective function value; **G** is the Euclidean norm of the final gradient; **PRELDF**, defined in (4.2.5), is the relative function reduction predicted for the last step attempted; **PRELDX**, defined in (4.2.6), is the relative size of the last

step.

The codes for stopping criteria in the column **C** have the following meaning. **A** means absolute function value convergence, i.e. (4.2.2) with $\epsilon_2 = 1.E-20$. This test criteria is only valid for those problems that have 0 as the optimal value; **R** means the relative function convergence, i.e. (4.2.3) with $\epsilon_3 = 1.E-10$; **X** means the relative iteration point convergence, i.e. (4.2.4) with $\epsilon_4 = 1.E-10$; **B** means that both conditions **R** and **X** hold. From the two tables, we see that **CONTR** performed better than **SUMSL** for a variety of test functions, especially for **Brown and Dennis, Wood, Biggs Exp6, Extended Rosenbrock, Chebyquad, Variable Dimensional, Beale function** and **Gulf function**.

Table 1. Numerical results for **CONTR** with $D_k = I$, $\Delta_0 = 1$.

PROBLEM	LS	N	NF	NG	C	F	G	PRELDF	PRELDX	
POWELL BADLY SCALED	0	2	200	108	X	0.1663E-07	.115E-06	.642E-08	.103E-10	1
ROSENBROCK	0	2	26	22	A	0.3530E-23	.122E-07	.100E+01	.445E-09	
ROSENBROCK	1	2	86	65	A	0.3672E-20	.529E-05	.100E+01	.142E-05	
ROSENBROCK	2	2	161	133	A	0.7309E-21	.295E-05	.100E+01	.586E-07	
BROWN BADLY SCALED	0	2	53	12	A	0.2133E-21	.823E-03	.100E+01	.135E-14	
BROWN BADLY SCALED	1	2	70	25	A	0.2118E-21	.413E-01	.100E+01	.414E-15	
BROWN BADLY SCALED	2	2	54	27	A	0.2213E-27	.366E-04	.100E+01	.815E-15	
BOX 3-D	0	3	35	30	A	0.1063E-22	.420E-09	.982E+00	.124E-09	
GAUSSIAN	0	3	7	5	B	0.1128E-07	.120E-07	.869E-09	.799E-10	
EXTENDED POWELL	0	4	71	61	A	0.7305E-20	.213E-08	.851E+00	.245E+00	
EXTENDED POWELL	1	4	65	61	A	0.9018E-20	.743E-10	.805E-01	.674E-01	
EXTENDED POWELL	2	4	71	62	A	0.7154E-20	.131E-08	.358E+00	.132E-01	
HELICAL VALLEY	0	3	32	23	X	0.5654E-18	.135E-05	.101E+01	.362E-10	
HELICAL VALLEY	1	3	41	32	A	0.4316E-21	.777E-07	.100E+01	.915E-09	
HELICAL VALLEY	2	3	40	27	A	0.6187E-22	.265E-07	.993E+00	.894E-10	
BROWN AND DENNIS	0	4	37	21	R	0.8582E+05	.779E-01	.432E-10	.186E-05	
BROWN AND DENNIS	1	4	49	31	R	0.8582E+05	.455E+00	.496E-10	.201E-05	
BROWN AND DENNIS	2	4	53	32	R	0.8582E+05	.350E+00	.658E-10	.790E-06	
WOOD	0	4	41	30	X	0.3336E-18	.312E-06	.105E+01	.143E-09	
WOOD	1	4	62	52	A	0.3604E-24	.429E-08	.999E+00	.189E-09	
WOOD	2	4	80	56	A	0.5847E-21	.809E-07	.992E+00	.322E-09	
BIGGS EXP6	0	6	40	35	R	0.5656E-02	.512E-07	.304E-12	.351E-07	4
PENALTY I	0	10	61	55	R	0.7088E-04	.910E-07	.657E-10	.982E-05	
PENALTY I	1	10	101	45	R	0.9284E-04	.625E-04	.138E-10	.770E-07	1
PENALTY I	2	10	302	199	R	0.7559E-04	.349E-04	.198E-11	.127E-05	1
EXTENDED ROSENBROCK	0	10	110	95	A	0.5679E-19	.425E-07	.980E+00	.226E-08	
EXTENDED ROSENBROCK	1	10	268	188	X	0.4266E-12	.358E-06	.784E-07	.148E-09	
EXTENDED ROSENBROCK	2	10	401	256	B	0.1469E-10	.234E-06	.397E-15	.188E-09	
CHEBYQUAD	0	7	24	15	X	0.4564E-16	.181E-06	.172E+01	.284E-09	
CHEBYQUAD	0	8	27	22	R	0.3517E-02	.330E-05	.956E-10	.899E-07	
CHEBYQUAD	0	9	47	34	A	0.2384E-19	.990E-08	.922E+00	.108E-08	
CHEBYQUAD	0	10	40	30	R	0.6504E-02	.107E-05	.444E-10	.185E-06	
PENALTY II	0	1	7	5	B	0.4894E+00	.704E-09	.919E-19	.807E-10	
PENALTY II	1	1	9	8	B	0.4894E+00	.234E-07	.102E-15	.268E-09	
PENALTY II	0	1	16	11	B	0.4894E+00	.338E-10	.107E-09	.358E-11	
PENALTY II	0	4	508	305	R	0.9378E-05	.123E-06	.275E-10	.188E-06	
PENALTY II	1	4	1201	768	R	0.9378E-05	.288E-04	.236E-09	.384E-05	
PENALTY II	2	4	641	507	R	0.9378E-05	.548E-05	.490E-10	.131E-04	
PENALTY II	0	10	174	104	R	0.2937E-03	.595E-06	.221E-10	.108E-04	
PENALTY II	1	10	448	316	R	0.2977E-03	.801E-04	.199E-10	.527E-05	1
PENALTY II	2	10	384	265	R	0.2965E-03	.400E-03	.415E-10	.554E-05	1
VARIABLE DIMENSIONED	0	10	19	15	A	0.5604E-21	.710E-06	.100E+01	.208E-09	
VARIABLE DIMENSIONED	1	10	29	25	X	0.4651E-17	.121E-08	.893E+01	.101E-09	
VARIABLE DIMENSIONED	2	10	45	32	X	0.7201E-17	.402E-06	.189E+01	.152E-09	
BEALE	0	2	17	15	A	0.7063E-23	.381E-08	.996E+00	.228E-09	
BEALE	1	2	94	65	A	0.3796E-20	.509E-07	.997E+00	.851E-08	2
BEALE	2	2	24	17	B	0.4441E+00	.225E-02	.312E-13	.170E-11	3
TRIGONOMETRIC	0	10	27	26	R	0.2795E-04	.204E-07	.105E-10	.123E-06	
TRIGONOMETRIC	1	10	47	31	A	0.3290E-21	.577E-09	.964E+00	.751E-09	2
TRIGONOMETRIC	2	10	54	43	R	0.4474E-06	.198E-07	.579E-09	.190E-08	
GULF	0	3	48	43	A	0.1001E-20	.140E-09	.166E-08	.983E+00	5
WATSON	0	9	23	10	B	0.2198E+02	.611E+02	.580E-10	.253E-10	
WATSON	0	12	16	2	B	0.2442E+02	.886E+02	.203E-10	.393E-11	

Table 2. Numerical results for **SUMSL** with $D_k = I$, $\Delta_0 = 1$.

PROBLEM	LS	N	NF	NG	C	F	G	PRELDF	PRELDX	
POWELL BADLY SCALED	0	2	219	160	A	0.25190E-23	.112E-06	.523E-07	.101E+01	1
ROSENBROCK	0	2	26	22	A	0.78857E-20	.381E-08	.223E-07	.100E+01	
ROSENBROCK	1	2	84	69	A	0.38627E-25	.823E-11	.218E-09	.100E+01	
ROSENBROCK	2	2	197	149	A	0.10314E-20	.142E-08	.236E-07	.999E+00	
BROWN BADLY SCALED	0	2	68	14	R	0.38169E-11	.210E+01	.810E-10	.100E+01	
BROWN BADLY SCALED	1	2	77	22	A	0.42077E-21	.410E-04	.945E-16	.100E+01	
BROWN BADLY SCALED	2	2	52	27	R	0.80529E-13	.489E+00	.307E-10	.100E+01	
BOX 3-D	0	3	37	31	A	0.17485E-20	.389E-10	.103E-07	.998E+00	
GAUSSIAN	0	3	7	6	R	0.11279E-07	.150E-08	.354E-09	.172E-07	
EXTENDED POWELL	0	4	79	75	A	0.56277E-20	.196E-11	.845E-01	.328E+00	
EXTENDED POWELL	1	4	70	64	A	0.35090E-20	.397E-14	.140E+00	.490E+00	
EXTENDED POWELL	2	4	71	68	A	0.46823E-20	.139E-13	.286E+00	.490E+00	
HELICAL VALLEY	0	3	31	26	A	0.39067E-19	.293E-08	.668E-08	.993E+00	
HELICAL VALLEY	1	3	40	28	R	0.21433E-18	.362E-08	.354E-09	.102E+01	
HELICAL VALLEY	2	3	42	29	A	0.29091E-21	.633E-09	.242E-10	.985E+00	
BROWN AND DENNIS	0	4	48	37	X	0.85822E+05	.651E-04	.246E-07	.100E-13	
BROWN AND DENNIS	1	4	37	29	X	0.85822E+05	.172E-01	.276E-05	.930E-10	
BROWN AND DENNIS	2	4	52	42	X	0.85822E+05	.343E-01	.196E-05	.625E-10	
WOOD	0	4	80	64	R	0.11859E-18	.101E-07	.220E-09	.102E+01	
WOOD	1	4	100	85	R	0.23124E-18	.134E-07	.632E-09	.100E+01	
WOOD	2	4	89	72	A	0.23038E-19	.332E-08	.156E-08	.101E+01	
BIGGS EXP6	0	6	43	42	X	0.56556E-02	.152E-07	.932E-06	.809E-10	4
PENALTY I	0	10	66	56	X	0.70877E-04	.145E-09	.861E-06	.601E-12	
PENALTY I	1	10	109	97	X	0.70877E-04	.118E-08	.393E-05	.169E-10	1
PENALTY I	2	10	358	279	X	0.70877E-04	.103E-08	.305E-05	.917E-11	1
EXTENDED ROSENBROCK	0	10	132	95	R	0.33067E-17	.436E-07	.849E-09	.101E+01	
EXTENDED ROSENBROCK	1	10	290	217	R	0.11515E-16	.100E-06	.119E-09	.906E+00	
EXTENDED ROSENBROCK	2	10	406	298	R	0.10263E-17	.158E-07	.540E-09	.959E+00	
CHEBYQUAD	0	7	30	21	R	0.92747E-15	.198E-06	.126E-09	.161E+01	
CHEBYQUAD	0	8	41	29	B	0.35169E-02	.134E-05	.954E-09	.287E-10	
CHEBYQUAD	0	9	46	36	R	0.31438E-18	.496E-09	.896E-09	.946E+00	
CHEBYQUAD	0	10	48	36	B	0.65040E-02	.344E-06	.129E-09	.602E-10	
PENALTY II	0	1	9	7	X	0.48940E+00	.809E-09	.506E-06	.362E-11	
PENALTY II	1	1	12	12	X	0.48940E+00	.147E-11	.870E-08	.106E-14	
PENALTY II	2	1	23	22	X	0.48940E+00	.122E-10	.322E-07	.147E-13	
PENALTY II	0	4	443	331	X	0.93763E-05	.781E-08	.520E-05	.298E-10	
PENALTY II	1	4	1325	996	X	0.93763E-05	.217E-07	.461E-05	.291E-10	
PENALTY II	2	4	820	636	X	0.93763E-05	.158E-07	.464E-05	.247E-10	
PENALTY II	0	10	208	173	X	0.29366E-03	.236E-06	.741E-05	.483E-10	
PENALTY II	1	10	653	532	X	0.29366E-03	.530E-06	.768E-05	.385E-10	1
PENALTY II	2	10	1800	1379	X	0.29376E-03	.123E-04	.236E-03	.602E-09	1
VARIABLE DIMENSIONED	0	10	21	21	A	0.20960E-21	.337E-10	.828E-07	.100E+01	
VARIABLE DIMENSIONED	1	10	25	25	A	0.18874E-21	.117E-09	.172E-07	.100E+01	
VARIABLE DIMENSIONED	2	10	39	37	A	0.71046E-25	.203E-11	.416E-10	.100E+01	
BEALE	0	2	17	16	A	0.35945E-26	.566E-12	.427E-10	.999E+00	
BEALE	1	2	21	15	B	0.39199E+00	.969E-02	.457E-09	.731E-09	2
BEALE	2	2	33	32	B	0.44786E+00	.323E-02	.112E-09	.105E-10	3
TRIGONOMETRIC	0	10	28	28	X	0.27951E-04	.128E-07	.968E-07	.757E-10	
TRIGONOMETRIC	1	10	40	38	X	0.87882E-03	.171E-07	.114E-06	.651E-11	2
TRIGONOMETRIC	2	10	36	32	R	0.26492E-15	.325E-07	.651E-09	.957E+00	
GULF	0	3	2	2	B	0.14000E-02	.000E+00	.000E+00	.000E+00	5
WATSON	0	9	48	6	B	0.21979E+02	.347E+02	.166E-13	.655E-13	
WATSON	0	12	58	3	B	0.24417E+02	.885E+02	.202E-13	.104E-14	

Notes to Tables 1-2.

1. **CONTR** reached the minimum but not as close as **SUMSL** did;
2. **CONTR** reached the minimum, but **SUMSL** failed;
3. Both **CONTR** and **SUMSL** failed to reach the minimum close enough;
4. **CONTR** and **SUMSL** both report convergence to a saddle point of **BIGGS EXP6**;
5. **CONTR** reached the minimum, but **SUMSL** reached only a saddle point.

CHAPTER 5

CONVERGENCE PROPERTIES AND CONCLUSIONS

In this chapter, we prove Q-superlinear convergence of the specific algorithm developed in the last chapter under reasonable conditions on the objective function and discuss some of the restrictions on the algorithm which can be expected during execution. We also make a short conclusion about the work in this part of the thesis.

5.1 Convergence Properties

Convergence results are very important, since to a certain extent, the rate of convergence of a method is as important as the fact that it converges. If it converges very slowly, we may never be able to see it converge. Therefore in this section, we outline some results which give insight into the rate of convergence. In particular we emphasize the notion of Q-superlinear convergence under certain general conditions on the objective function and discuss some of the restrictions which can be expected during implementation.

A reasonable algorithm should at least be Q-linearly convergent in the sense that if $\{x_k\}$ is generated by the algorithm and $\{x_k\}$ converges to x^*, then for some norm $\|\cdot\|$, there is a constant $\alpha \in (0,1)$ and $k_0 > 0$ such that

$$\|x_{k+1} - x^*\| \leq \alpha \|x_k - x^*\|, \tag{5.1.1}$$

for all $k > k_0$. This guarantees that eventually the error will be decreased by a factor $\alpha < 1$. To be competitive, a nonlinear minimization algorithm should

be Q-superlinearly convergent in the sense that there exists $\{\alpha_k\}$, converging to zero, such that (5.1.1) holds for α_k instead of α.

Let's denote $B_k = C_k^{-1}$ in *Algorithm 4.1.3*. For the purpose of simplicity, we choose D_k to be the unit matrix. Then (3.6) becomes

$$\begin{cases} \min & f_k + g_k^T v + \frac{1}{2} v^T B_k v, \\ & s = v/(1 + \alpha_k g_k^T v), \quad \|s\| \le \Delta_k. \end{cases} \tag{5.1.2}$$

We denote by $\varphi_k(s)$ the objective function in (5.1.2) in terms of the variable s.

LEMMA 5.1.1. *From the implementation of the Algorithm 4.1.3, we have*

$$f_k - \varphi_k(s_k) \ge \frac{1}{2}\|g_k\| \min\{ \frac{\|g_k\|}{\|B_k\|}, \quad \frac{\Delta_k}{1 + \alpha_k \Delta_k \|g_k\|} \}. \tag{5.1.3}$$

PROOF: Let $v(\tau) = -\tau g_k$ for $\tau > 0$ and correspondingly $s(\tau) = -\tau g_k/(1 - \alpha_k \tau \|g_k\|^2)$. The minimum of

$$\varphi(s(\tau)) = f_k - \tau g_k^T g_k + \frac{1}{2}\tau^2 g_k^T B_k g_k$$

is reached at $\tau^* = g_k^T g_k / g_k^T B_k g_k$. Now if $\|s(\tau^*)\| = \tau^* \|g_k\|/|1 - \tau^* \alpha_k \|g_k\|^2| \le \Delta_k$, then

$$\varphi(s_k)) \le \varphi(s(\tau^*)) \le f_k - \frac{1}{2}\frac{(g_k^T g_k)^2}{g_k^T B_k g_k} \le f_k - \frac{1}{2}\frac{\|g_k\|^2}{\|B_k\|}. \tag{5.1.4}$$

Otherwise $\|s(\tau^*)\| = \tau^* \|g_k\|/|1 - \tau^* \alpha_k \|g_k\|^2| > \Delta_k$. We choose τ_0 such that $\tau_0 \|g_k\|/|1 - \tau_0 \alpha_k \|g_k\|^2| = \Delta_k$, i.e. $\tau_0 = \Delta_k/|\|g_k\| + \alpha_k \Delta_k \|g_k\|^2|$. Then $\tau_* > \tau_0$, i.e.

$$\frac{g_k^T g_k}{g_k^T B_k g_k} > \frac{\Delta_k}{\|g_k\| + \alpha_k \Delta_k g_k^T g_k}.$$

Therefore

$$
\begin{aligned}
& -(f_k - \varphi_k(s_k)) \le -f_k + \varphi_k(s(\tau_0)) \\
\le & -\frac{\Delta_k \|g_k\|^2}{\|g_k\| + \alpha_k \Delta_k \|g_k\|^2} + \frac{1}{2} \frac{\Delta_k^2 g_k^T B_k g_k}{(\|g_k\| + \alpha_k \Delta_k g_k^T g_k)^2} \\
= & \frac{-\Delta_k}{\|g_k\| + \alpha_k \Delta_k g_k^T g_k} \left(g_k^T g_k - \frac{1}{2} \frac{\Delta_k g_k^T B_k g_k}{\|g_k\| + \alpha_k \Delta_k g_k^T g_k} \right) \\
= & \frac{-\Delta_k g_k^T B_k g_k}{\|g_k\| + \alpha_k \Delta_k g_k^T g_k} \left(\frac{g_k^T g_k}{g_k^T B_k g_k} - \frac{1}{2} \frac{\Delta_k}{\|g_k\| + \alpha_k \Delta_k g_k^T g_k} \right) \\
\le & -\frac{1}{2} \frac{\Delta_k g_k^T B_k g_k}{\|g_k\|(1 + \alpha_k \Delta_k \|g_k\|)} \frac{g_k^T g_k}{g_k^T B_k g_k} \\
= & -\frac{1}{2} \frac{\Delta_k \|g_k\|}{1 + \alpha_k \Delta_k \|g_k\|},
\end{aligned}
\tag{5.1.5}
$$

which, together with (5.1.4), proves the Lemma. □

This lemma shows that each iteration has a sufficient decrease in terms of model approximation $f_k - \varphi_k(s_k)$.

LEMMA 5.1.2. *Let* $f(x) : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable and* $\Omega(x_0) := \{ x : f(x) \le f(x_0) \}$ *be bounded. Then* $\{\|g_k\|\}$ *is bounded by some real number* b . *Furthermore, from Algorithm 4.1.3, if* $\|s_k\| \to 0$, *there exists an integer* k_1 *s.t.*

$$
c \le 1 - \alpha_k g_k^T s_k \le c_1, \tag{5.1.6}
$$

for all $k \ge k_1$ *and some positive constants* c , c_1 .

PROOF: The boundedness of $\{\|g_k\|\}$ follows from the continuous differentiability of $f(x)$ together with the boundedness of $\Omega(x_0)$.

From (4.1.7), $\alpha_k = (1-\nu)/\nu g_k^T s_k$, or $\alpha_k = sign(\frac{1-\nu}{\nu g_k^T s_k}) \frac{\alpha_0}{\Delta_k \|g_k\|}$, or $\alpha_k = sign(\frac{1-\nu}{\nu g_k^T s_k}) \tilde{\alpha}$. In the first case, we have

$$
1 - \alpha_k g_k^T s_k = 2 - \frac{1}{\nu} \in [2 - \frac{1}{\nu_1}, 2 - \frac{1}{\nu_2}] \in (0, \frac{7}{4}), \tag{5.1.7}
$$

by the choice of ν_1 , ν_2 from *Algorithm 4.1.3*.

In the second case, we have

$$|\frac{1-\nu}{\nu g_k^T s_k}| \geq \frac{\alpha_0}{\Delta_k \|g_k\|}.$$

Therefore

$$\begin{aligned} 1-\alpha_k g_k^T s_k &= 1 - sign(\frac{1-\nu}{\nu g_k^T s_k})\frac{\alpha_0}{\Delta_k\|g_k\|} g_k^T s_k \\ &= 1 - \frac{1-\nu}{\nu g_k^T s_k}|\frac{\nu g_k^T s_k}{1-\nu}|\frac{\alpha_0}{\Delta_k\|g_k\|} g_k^T s_k \\ &= 1 - \frac{1-\nu}{\nu}|\frac{\nu g_k^T s_k}{1-\nu}|\frac{\alpha_0}{\Delta_k\|g_k\|}. \end{aligned} \tag{5.1.8}$$

If $\frac{1-\nu}{\nu} < 0$, then from (5.1.7),

$$1-\alpha_k g_k^T s_k \leq 1 - \frac{1-\nu}{\nu}\frac{\Delta_k\|g_k\|}{\alpha_0}\frac{\alpha_0}{\Delta_k\|g_k\|} = 1 - \frac{1-\nu}{\nu} \leq 2 - \frac{1}{\nu_2}, \tag{5.1.9}$$

$$1-\alpha_k g_k^T s_k \geq 1 - \frac{\alpha_0}{\Delta_k\|g_k\|}|g_k^T s_k| \geq 1 - \frac{\alpha_0}{\Delta_k\|g_k\|}\|g_k\|\|s_k\| \geq 1-\alpha_0. \tag{5.1.10}$$

If $\frac{1-\nu}{\nu} > 0$, then from (5.1.7),

$$1-\alpha_k g_k^T s_k \geq 1 - \frac{1-\nu}{\nu}\frac{\Delta_k\|g_k\|}{\alpha_0}\frac{\alpha_0}{\Delta_k\|g_k\|} = 1 - \frac{1-\nu}{\nu} \geq 2 - \frac{1}{\nu_1} > 0, \tag{5.1.11}$$

$$1-\alpha_k g_k^T s_k \leq 1 + \frac{\alpha_0}{\Delta_k\|g_k\|}|g_k^T s_k| \leq 1 + \frac{\alpha_0}{\Delta_k\|g_k\|}\|g_k\|\|s_k\| \leq 1+\alpha_0. \tag{5.1.12}$$

In the third case, we have

$$1-\alpha_k g_k^T s_k \leq 1 + \tilde{\alpha} b \overline{\Delta}, \tag{5.1.13}$$

$$1-\alpha_k g_k^T s_k \geq 1 - \tilde{\alpha}\|g_k\|\|s_k\| \geq 1 - \tilde{\alpha} b\|s_k\| \geq 1-\alpha_0, \tag{5.1.14}$$

as $\|s_k\| \to 0$, there exists positive integer k_1 s.t. $\|s_k\| \leq \alpha_0/\tilde{\alpha} b$ for all $k \geq k_1$.

The Lemma is proved if we choose

$$c = \min\{\, 2 - \frac{1}{\nu_1},\, 1-\alpha_0,\, \}, \tag{5.1.15}$$

$$c_1 = \max\{\, 2 - \frac{1}{\nu_2},\, 1+\alpha_0,\, 1+\tilde{\alpha} b\overline{\Delta}\,\}. \tag{5.1.16}$$

Here $c > 0$, as $\alpha_0 < 1$ and $\nu_1 > \frac{1}{2}$ from *Algorithm 4.1.3*. □

As we will see, $1 - \alpha_k g_k^T s_k$ is an important denominator factor. The boundedness of this factor is essential for the proof of the next lemma.

LEMMA 5.1.3. *Let* $f(x) : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable,* $\Omega(x_0) := \{ x : f(x) \leq f(x_0) \}$ *be bounded, and* $g(x) = \nabla f(x)$ *be Lipschitz continuous. Suppose* $\{x_k\}$ *is the sequence produced by Algorithm 4.1.3 applied to* $f(x)$ *for a given starting point* x_0. *Then under the bounded deterioration condition of* B_k, *i.e.*

$$\|B_k\| \leq m_1 + m_2 \sum_{i=1}^{k} \|s_i\|, \tag{5.1.17}$$

there is no constant $\epsilon > 0$ *such that* $\|g_k\| \geq \epsilon$ *for all* k.

PROOF: Assume there is $\epsilon > 0$ such that $\|g_k\| \geq \epsilon$ for all k. From *Lemma 5.1.1*,

$$\begin{aligned} f_k - \varphi_k(s_k) \geq & \frac{1}{2}\|g_k\| \min\Big\{ \frac{\|g_k\|}{\|B_k\|}, \quad \frac{\Delta_k}{1 + \alpha_k \Delta_k \|g_k\|} \Big\} \\ \geq & \frac{1}{2}\|g_k\| \frac{\Delta_k \|g_k\|}{\Delta_k \|B_k\| + \|g_k\|(1 + \alpha_k \Delta_k \|g_k\|)}, \end{aligned} \tag{5.1.18}$$

because $\min\{|a|, |b|\} \geq |ab|/(|a| + |b|)$ for any real numbers a, b. Then from **Step 4** of *Algorithm 4.1.3*,

$$f_k - f_{k+1} \geq \frac{1}{2}\eta_1 \|g_k\| \frac{\Delta_k \|g_k\|}{\Delta_k \|B_k\| + \|g_k\|(1 + \alpha_k \Delta_k \|g_k\|)}. \tag{5.1.19}$$

Since $f(x)$ is bounded below and $f_{k+1} < f_k$ for all k, then $\sum_{k=1}^{\infty}(f_k - f_{k+1})$ is convergent, i.e.

$$\sum_{k=1}^{\infty} \frac{\Delta_k \|g_k\|^2}{\Delta_k \|B_k\| + \|g_k\|(1 + \alpha_k \Delta_k \|g_k\|)} \tag{5.1.20}$$

converges. By the bounded deterioration condition and $\Delta_k \leq \overline{\Delta}$, we know

$$\sum_{k=1}^{\infty} \frac{\Delta_k}{\overline{\Delta}(\alpha_1 + \alpha_2 \sum_{i=1}^{k} \Delta_i)/\epsilon^2 + (1 + \alpha_k \Delta_k \|g_k\|)/\epsilon}$$

is finite. And by the stability requirement (4.1.7) of *Algorithm 4.1.3*,

$$0 < 1 - \alpha_0 \leq 1 + \alpha_k \Delta_k \|g_k\| \leq 1 + \alpha_0.$$

We conclude that $\sum_{k=1}^{\infty} \Delta_k$ converges, as for any positive series $\{a_k\}$, the convergence of

$$\sum_{k=1}^{\infty} a_k / (\sum_{i=1}^{k} a_i + E)$$

for some constant E implies the convergence of $\sum_{k=1}^{\infty} a_k$, see [Flet72]. Therefore $\lim \Delta_k = 0$ and $\{\|B_k\|\}$ is bounded. Assume, without loss of generality, that $\|B_k\| \leq b$.

As $\{\alpha_k\}$ and $\{\|g_k\|\}$ are bounded from *Algorithm 4.1.3* and *Lemma 5.1.2*, there exists a positive integer $k_2 \geq k_1$, which is large enough such that

$$\Delta_k / (1 + \alpha_k \Delta_k \|g_k\|) \leq \epsilon / b, \quad \text{for all} \quad k \geq k_2.$$

Then from (5.1.18), we have

$$\begin{aligned} f_k - \varphi_k(s_k) &\geq \frac{1}{2}\|g_k\| \frac{\Delta_k}{\Delta_k \|B_k\| / \|g_k\| + (1 + \alpha_k \Delta_k \|g_k\|)} \\ &\geq \frac{1}{2}\|g_k\| \frac{\Delta_k}{\overline{\Delta} b / \epsilon + (1 + \alpha_0)} \geq \frac{1}{2}\beta_1 \|g_k\| \|s_k\|, \end{aligned} \tag{5.1.21}$$

where $\beta_1 = 1/(\overline{\Delta} b/\epsilon + 1 + \alpha_0)$. So we have

$$\begin{aligned} &\left|1 + \frac{g_k^T v_k}{f_k - \varphi_k(s_k)}\right| = \frac{1}{2} \frac{|v_k^T B_k v_k|}{|f_k - \varphi_k(s_k)|} \\ \leq &\frac{1}{2} \frac{1}{(1 - \alpha_k s_k^T g_k)^2} \frac{\|s_k\|^2 \|B_k\|}{\frac{1}{2}\beta_1 \|g_k\| \|s_k\|} \\ \leq &\frac{b}{c^2 \beta_1 \epsilon} \Delta_k, \end{aligned} \tag{5.1.22}$$

from *Lemma 5.1.2*.

From the Lipschitzian continuity of $g(x)$ and the Mean-Value Theorem,

$$
\begin{aligned}
&|\frac{f_k - f_{k+1}}{f_k - \varphi_k(s_k)} - 1| = |\frac{f_{k+1} - \varphi_k(s_k)}{f_k - \varphi_k(s_k)}| \\
\leq &|\frac{f_{k+1} - f_k - g_k^T v_k}{f_k - \varphi_k(s_k)}| + |\frac{v_k^T B_k v_k}{f_k - \varphi_k(s_k)}| \\
\leq &|\frac{g(x_k + \theta s_k)^T s_k - g_k^T s_k/(1 - \alpha_k g_k^T s_k)}{|f_k - \varphi_k(s_k)|}| + |\frac{v_k^T B_k v_k}{|f_k - \varphi_k(s_k)|}| \\
\leq &|\frac{(g(x_k + \theta s_k) - g(x_k))^T s_k + (1 - 1/(1 - \alpha_k g_k^T s_k)) g_k^T s_k}{|f_k - \varphi_k(s_k)|}| \\
&\quad + |\frac{v_k^T B_k v_k}{|f_k - \varphi_k(s_k)|}| \\
\leq &\frac{\theta L \|s_k\|^2 + |\alpha_k| |g_k^T s_k|^2/|1 - \alpha_k g_k^T s_k|}{|f_k - \varphi_k(s_k)|} + |\frac{v_k^T B_k v_k}{|f_k - \varphi_k(s_k)|}| \\
\leq &\frac{L\|s_k\|^2}{\frac{1}{2}\beta_1 \|g_k\| \|s_k\|} + \frac{(\tilde{\alpha} b^2/c)\|s_k\|^2}{\frac{1}{2}\beta_1 \|g_k\| \|s_k\|} + \frac{b}{c^2 \beta_1 \epsilon} \Delta_k, \\
\leq &\beta \Delta_k, \qquad \text{for all} \quad k \geq k_2,
\end{aligned}
\tag{5.1.23}
$$

for Lipschitzian constant L, $\theta \in (0,1)$ and some $\beta > 0$.

So by the updating of Δ_k in **step 4** c), d) of the *Algorithm 4.1.3*, (5.1.23) implies that Δ_k is not going to 0, which is a contradiction. □

Lemma 5.1.3 proves that at least one subsequence of $\{x_k\}$ converges to a critical point of $f(x)$. Therefore if $\{x_k\}$ converges, then it must converge to the critical point.

COROLLARY 5.1.4. *Let the hypotheses of Lemma 5.1.3 be satisfied. If* $\{x_k\}$ *converges to* x^*, *then* $g(x^*) = 0$.

THEOREM 5.1.5. *Under the assumptions of Lemma 5.1.3, if* $\{\|B_k\|\}$ *is bounded, then we have* $\lim \|g_k\| = 0$.

PROOF: Let $\|B_k\| \leq b$ for all positive integer k, suppose there is a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ such that $\|g(x_{k_j})\| \geq \epsilon > 0$, for all $j = 1, 2, \cdots$. By *Lemma 5.1.3*, on the other hand, there is another subsequence $\{x_{k'_j}\}$ of $\{x_k\}$

such that $\|g(x_{k'_j})\| \to 0$, as $j \to \infty$. Without loss of generality, we assume that $k'_j \in [k_j, k_{j+1}) := \{ k_j, k_j + 1, \cdots, k_{j+1} - 1 \}$ for all $j = 1, 2, \cdots$. Therefore we can select an integer l_j for each j such that

$$l_j = \max\{ l \in [k_j, k_{j+1}) : \|g_i\| \geq \frac{\epsilon}{2}, \quad \text{for all} \quad i \quad \text{s.t.} \quad k_j \leq i \leq l \}. \tag{5.1.24}$$

From (5.1.18) and **step 4** of *Algorithm 4.1.3*,

$$f_l - f_{l+1} \geq \frac{1}{4}\eta_1 \epsilon \min\{ \frac{\Delta_l}{\alpha_0 + 1}, \quad \frac{\epsilon}{2b} \} \qquad \text{for all} \quad l \in [k_j, l_j], \tag{5.1.25}$$

therefore

$$f_{k_j} - f_{l_j+1} \geq \frac{1}{4}\eta_1 \epsilon \min\{ \frac{1}{\alpha_0 + 1} \sum_{i=k_j}^{l_j} \Delta_i, \quad \frac{\epsilon}{2b} \} \qquad \text{for all} \quad l \in [k_j, l_j]. \tag{5.1.26}$$

But $f_{k_j} - f_{l_j+1} \to 0$, so $\sum_{i=k_j}^{l_j} \Delta_i \to 0$ as $j \to \infty$, which means

$$\|x_{k_j} - x_{l_j+1}\| \leq \sum_{i=k_j}^{l_j} \Delta_i \to 0, \qquad \text{as} \quad j \to \infty. \tag{5.1.27}$$

Let j be large enough such that $\|x_{k_j} - x_{l_j+1}| \leq \epsilon/2L$. Then by the Lipschitzian continuity of $g(x)$, we have

$$\|g_{k_j}\| \leq \|g_{k_j} - g_{l_j+1}\| + \|g_{l_j+1}\| \leq L\|x_{k_j} - x_{l_j+1}\| + \frac{\epsilon}{2} \leq \epsilon, \tag{5.1.28}$$

which is a contradiction. □

Theorem 5.1.5 proves the convergence property of the *Algorithm 4.1.3*. Under further assumptions, we can prove the Q-superlinear convergence property. At first, we assume that $f(x)$ is twice continuously differentiable in a neighborhood of x^* , a local minimum of $f(x)$. Then we have

$$g(x_k) = g(x_k) - g(x^*) = \int_0^1 \nabla^2 f(x^* + \theta(x_k - x^*))(x_k - x^*) d\theta, \tag{5.1.29}$$

$$f_k - f(x^*) = \int_0^1 (1 - \theta)(x_k - x^*)^T \nabla^2 f(x^* + \theta(x_k - x^*))(x_k - x^*) d\theta. \tag{5.1.30}$$

If $x_k \to x^*$ and $\nabla^2 f(x^*)$ is positive definite, then there exists real numbers $m > 0$, $M > 0$ and positive integer $k_3 > k_2$ such that

$$m\|x_k - x^*\| \le \|g_k\| \le M\|x_k - x^*\|, \qquad \text{for all} \quad k \ge k_3, \tag{5.1.31}$$

$$\frac{1}{2}m\|x_k - x^*\|^2 \le f_k - f(x^*) \le \frac{1}{2}M\|x_k - x^*\|^2, \qquad \text{for all} \quad k \ge k_3. \tag{5.1.32}$$

Therefore we have

$$\sqrt{f_k - f(x^*)} \le \sqrt{\frac{M}{2m^2}}\,\|g_k\|, \qquad \text{for all} \quad k \ge k_3, \tag{5.1.33}$$

$$\begin{aligned} \|s_k\| = \|x_{k+1} - x_k\| &\le \|x_k - x^*\| + \|x_{k+1} - x^*\| \\ &\le \|x_k - x^*\| + \sqrt{2(f_{k+1} - f(x^*))/m} \\ &\le \|x_k - x^*\| + \sqrt{2(f_k - f(x^*))/m} \\ &\le (1 + \sqrt{\frac{M}{m}})\frac{\|g_k\|}{m}, \qquad \text{for all} \quad k \ge k_3. \end{aligned} \tag{5.1.34}$$

THEOREM 5.1.6. *Suppose $f(x)$ is twice continuously differentiable, $\{x_k\}$ is the sequence generated from Algorithm 4.1.3 with initial point x_0 s. t. $\Omega(x_0)$ is bounded. Assume that $\{x_k\}$ converges to x^* with Hessian of $f(x)$ at x^* positive definite. If $\{B_k\}$ is bounded, and satisfies*

$$\lim_{k\to\infty} \frac{\|(B_k - \nabla^2 f(x^*))v_k\|}{\|v_k\|} = 0. \tag{5.1.35}$$

Then there exists positive integer k_3 s.t.

$$\frac{\|s_k\|}{\|g_k\|} \le c_2, \qquad \text{for all} \quad k \ge k_3, \tag{5.1.36}$$

and the rate of convergence of $\{x_k\}$ is Q-superlinear, i.e.

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0.$$

PROOF: (5.1.36) is immediate from (5.1.34) for $c_2 = \frac{1}{m}(1+\sqrt{\frac{M}{m}})$. From (5.1.3) and (5.1.36), whenever $k \geq k_3$,

$$f_k - \varphi_k(s_k) \geq \frac{1}{2}\|s_k\|\frac{\|g_k\|}{\|s_k\|}\min\{\frac{1}{\|B_k\|}\frac{\|g_k\|}{\|s_k\|}\|s_k\|,\ \frac{\Delta_k}{\alpha_0+1}\} \geq c_3\|s_k\|^2.$$

here $c_3 = \frac{1}{2}\frac{1}{c_2}\min\{\frac{1}{c_2 b}, \frac{1}{\alpha_0+1}\}$. By Taylor's expansion,

$$\begin{aligned}
&f_{k+1} - \varphi_k(s_k) \\
=&\frac{g_k^T v_k}{1+\alpha_k g_k^T v_k} + \frac{1}{2}\frac{1}{(1+\alpha_k g_k^T v_k)^2}v_k^T M_k v_k - g_k^T v_k - \frac{1}{2}v_k B_k v_k \\
=&\frac{-\alpha_k g_k^T v_k}{1+\alpha_k g_k^T v_k}g_k^T v_k - \frac{1}{2}v_k^T(B_k - \nabla^2 f(x^*))v_k \\
&-\frac{1}{2}v_k^T(\nabla^2 f(x^*) - M_k)v_k - \frac{1}{2}\Big(1 - \frac{1}{(1+\alpha_k g_k^T v_k)^2}\Big)v_k^T M_k v_k,
\end{aligned} \tag{5.1.37}$$

here $M_k = \nabla^2 f\big(x_k + \theta v_k/(1+\alpha_k g_k^T v_k)\big)$ for some $\theta \in (0,1)$. From *Theorem 5.1.5*, we have $\|g_k\| \to 0$. Because $v_k = s_k/(1-\alpha_k g_k^T s_k)$, and from *Lemma 5.1.2* and (5.1.35), we have

$$\frac{\|f_{k+1} - \varphi_k(s_k)\|}{\|s_k\|^2} \to 0 \quad \text{as} \quad k \to \infty. \tag{5.1.38}$$

Therefore

$$\lim_{k\to\infty}\Big|\frac{f_k - f_{k+1}}{f_k - \varphi_k(s_k)} - 1\Big| \leq \lim_{k\to\infty}\Big|\frac{f_{k+1} - \varphi_k(s_k)}{c_3\|s_k\|^2}\Big| = 0.$$

So there exists $k_4 \geq k_3$ such that $\Delta_k \geq \Delta_{k_4}$, for all $k \geq k_4$, i.e. $\{\Delta_k\}$ is bounded below. On the other hand, as $\|g_k\| \to 0$, when k is sufficiently large, $\|B^{-1}g_k/(1-\alpha_k g_k B_k^{-1}g_k)\| \leq \Delta_k$, and $v_k = -B_k^{-1}g_k$, $s_k = -B^{-1}g_k/(1-\alpha_k g_k B_k^{-1}g_k)$ holds, i.e. when k is sufficiently large, we have $g_{k+1} - g_k - B_k v_k = g_{k+1}$. By an analogous way to prove (5.1.38), we have

$$\frac{\|g_{k+1}\|}{\|v_k\|} = \frac{\|g_{k+1} - g_k - B_k v_k\|}{\|v_k\|},$$

which goes to 0 . Therefore by *Lemma 5.1.2* and (5.1.36),

$$\frac{\|g_{k+1}\|}{\|s_k\|} \to 0, \qquad \text{and} \quad \frac{\|g_{k+1}\|}{\|g_k\|} \to 0,$$

and by (5.1.31) we obtain

$$\frac{m\|x_{k+1} - x^*\|}{M\|x_k - x^*\|} \to 0,$$

which proves the Q-superlinear convergence of the method. □

Remark Assumption (5.1.35) is standard; e.g. in [DeMo74], this condition is used to prove Q-superlinear convergence for the quadratic model for both **DFP** and **BFGS** updates.

5.2 Conclusions

Conic model methods with trust region approach were proposed in this part of the thesis. We discussed its motivations and theoretical background, which involve necessary and sufficient conditions for solutions to the model trust region subproblems (i.e. *Theorem 3.3.1*, *Theorem 3.3.2* and *Theorem 3.3.3*). A specific computer program shows that the class of methods has very satisfactory numerical performance for most of the standard test functions. The convergence properties proved in §5.1 show that a typical algorithm has Q-superlinear rate of convergence. All the topics suggest that conic model methods with trust region approach are very competitive.

There is still quite a lot to be done in this area. We expect that the boundedness of B_k in §5.1 (i.e. C_k in §4.1) can be proved under reasonable conditions. Like the work done in [BSSh87] and [CGTo20] for quadratic models, further extension of this framework to constrained minimization is also interesting.

PART II

CALMNESS FOR OPTIMIZATION PROBLEMS WITH EQUALITY CONSTRAINTS

CHAPTER 6

CALMNESS AND COUNTEREXAMPLES

The notion of calmness was introduced by F.H. Clarke and R.T. Rockafellar and first appeared in [Clar76]. Since it first appeared, it has been recognized as a fundamental concept in nonsmooth analysis, optimization theory and sensitivity analysis. Subsequently many variations of the original definitions and applications have been proposed (see [Fiac83], [Rock85], [Ward90] and [Burk91]). In this chapter, we review some of its definitions and applications, including its attributive property as a constraint qualification, its equivalence to the exactness of certain penalty functions and an assertion that "*almost all problems are calm*" for inequality constrained problems. For optimization problems that incorporate equality constraints and an abstract constraint, we don't have the same assertion because counterexamples are provided in this chapter.

6.1 Calmness and Constraint Qualifications

For problem $\mathcal{P}$ defined in (1.1.1), the Lagrangian is the function $L(x, \lambda, r, v, k,) : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^s \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}$ defined by

$$L(x, \lambda, r, v, k) := \lambda f(x) + \langle r, g(x) \rangle + \langle v, h(x) \rangle + k|(\lambda, r, v)|d_C(x), \qquad (6.1.1)$$

where $d_C(x)$ denotes the usual distance function associated with the set C, and $\langle \cdot, \cdot \rangle$ denotes the usual dot product. The following Lagrange Multiplier Theorem, providing necessary conditions for local optimal points, is from [Clar76].

THEOREM 6.1.1. *Let x be a local solution to $\mathcal{P}$. Then for every k sufficiently large, there exist $\lambda \geq 0$, $r \geq 0$, and v not all zero such that $\langle r, g(x) \rangle = 0$ and*

$$0 \in \partial_x L(x, \lambda, r, v, k), \tag{6.1.2}$$

where $\partial_x L(x, \lambda, r, v, k)$ is the generalized gradient, defined by (1.1.3), of the function $L(x, \lambda, r, v, k)$ relative to the variable x.

The necessary condition can be viewed as degenerate when the multiplier λ corresponding to f vanishes (because the function f being minimized is not involved). Various conditions have been proposed in order that the multiplier rule holds in normal form (i.e. the multiplier rule holds for some $\lambda > 0$). Those conditions are called *constraint qualifications*. Most of the constraint qualifications are conditions that ensure that the set of Lagrange multipliers in the degenerate case is $\{0\}$, which means that the degenerate case doesn't happen, e.g., the *Mangasarian-Fromowitz condition*, and *Slater condition*. *Calmness* is another type of constraint qualification that ensures that the multiplier rule holds for some $\lambda > 0$. For Lipschitz problems, calmness is weaker than the *Mangasarian-Fromowitz condition* or *Slater condition* (see [Clar83]). For non-Lipschitz problems, this is not the case; an example is provided by D. Ward in [Ward90].

DEFINITION 6.1.2. *For $p \in \mathbb{R}^s$, $q \in \mathbb{R}^m$, we define the value function of problem $\mathcal{P}(p, q)$ to be*

$$V(p, q) := \inf\{ f(x) : x \text{ feasible for } \mathcal{P}(p, q) \} \tag{6.1.3}$$

if $\mathcal{P}(p, q)$ has a nonempty feasible set; otherwise $V(p, q) := +\infty$.

The value function is sometimes also called the *marginal function* (see [Rock81] or [Ward90]). The following result is from [Clar83].

PROPOSITION 6.1.3. *([Clar83], 6.4.2.) Let* $V(p,q)$ *be finite, and suppose that one has*

$$\liminf_{(\alpha,\beta)\to(0,0)} \frac{V(p+\alpha,q+\beta)-V(p,q)}{|(\alpha,\beta)|} > -\infty. \tag{6.1.4}$$

Then for any (local) solution x *to* $\mathcal{P}(p,q)$, $\mathcal{P}(p,q)$ *is calm at* x *; see (1.1.4).*

DEFINITION 6.1.4. *We call a problem* $\mathcal{P}(p,q)$ *calm if* $V(p,q)$ *is finite and (6.1.4) holds.*

So if $\mathcal{P}(p,q)$ is calm, then it is calm at any local minimizer. But the converse is not true for general functions. An example can be found in [GaDu84].

PROPOSITION 6.1.5. *([Clar83], 6.4.4.) If* $\mathcal{P}$ *is calm, then the multiplier rule, Theorem 1.1, holds with* $\lambda = 1$.

From this proposition we know that calmness has the attribute of a constraint qualification, i.e. calmness ensures that the multiplier rule holds for some $\lambda > 0$.

6.2 Calmness and its Other Properties

J.V. Burke used another definition of calmness, saying that $\mathcal{P}$ is calm at x for modulus M and radius ϵ if for all $x' \in x + \epsilon\mathcal{B}$ feasible for $\mathcal{P}(p,q)$, (1.1.6) holds. So Burke's definition of calmness is stronger than that of Clarke. But under continuity conditions on g and h , it is proved that the two definitions are equivalent.

PROPOSITION 6.2.1. *([Burk91]) Assume* f , g , *and* h *are Lipschitz continuous. Let* $\bar{x} \in \mathbb{R}^n$, $p \in \mathbb{R}^s$, $q \in \mathbb{R}^m$ *such that* $\bar{x}$ *is feasible for* $\mathcal{P}(p,q)$. $\mathcal{P}(p,q)$ *is calm at* $\bar{x}$ *with modulus* M' *and radius* ϵ *iff the following condition holds:*

there exists a positive number M, *for any* $x' \in \mathbb{R}^n$, $\alpha \in \mathbb{R}^s$, $\beta \in \mathbb{R}^m$ *such that* x' *is feasible for* $\mathcal{P}(p+\alpha, q+\beta)$, *one has*

$$f(x') - f(\bar{x}) + M|(\alpha, \beta)| \geq 0. \tag{6.2.1}$$

The following definitions are from [Ward90].

DEFINITION 6.2.2. *Let* f *be an extended valued function on* $\mathbb{R}^n$, *i.e. a function from* $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. *Given a subset* S *of* $\mathbb{R}^n$, *we say that* $\bar{x} \in S$ *is an* ϵ*-minimum of* f *on* S *for some* $\epsilon \geq 0$ *if* $f(\bar{x}) \leq f(x)$ *for all* $x \in S \cap (\bar{x} + \epsilon\mathcal{B})$.

DEFINITION 6.2.3. *Let* x_0 *be a local minimum of* $\mathcal{P}(p,q)$. *Define*

$$\big(g(x)\big)_i^+ := \max\{\, g_i(x), 0 \,\}. \tag{6.2.2}$$

For $M > 0$, *the penalty function*

$$\phi_M(x) := f(x) + M|\big((g(x) - p)^+, h(x) - q\big)| \tag{6.2.3}$$

is said to be exact at x_0 *on a set* $S \subset \mathbb{R}^n$ *if* x_0 *is a local minimizer for the problem*

$$\min\{\, \phi_M(x) : x \in S \,\}.$$

Hereafter we denote by $\mathcal{P}_1(p)$ the special case of the problem $\mathcal{P}(p,q)$ without equality constraints, by $\mathcal{P}(q)$ the special case of the problem $\mathcal{P}(p,q)$ without inequality constraints.

The term *exact penalization* refers to a reduction principle in constrained optimization wherein one replaces a constrained minimization problem by an unconstrained minimization problem where the objective function is finite everywhere and which possesses a common local minimum under various conditions. The

reason for the replacement is largely from the numerical point of view; that is how we solve a constrained minimization problem by solving sequential unconstrained minimization problems numerically, i.e. penalty methods mentioned in §1.2. The particular reduction discussed herein originated with Eremin in [Erem66] and Zangwill in [Zang67] (see also [Piet69]). Recently, D. Ward established that, for non-Lipschitz functions, if f is lower semicontinuous, the calmness of $\mathcal{P}_1$ at $\bar{x}$ is equivalent to the exactness of certain penalty function at $\bar{x}$ (see [Ward90]). In [Burk91], Burke also established an equivalence between the notion of calmness and the viability of the Eremin-Zangwill penalization procedure for set valued constraints. Their work can be specified for program $\mathcal{P}(p,q)$ as the following theorem. The necessity of the theorem has already appeared in [Clar83], and a simpler proof is provided with the help of *Proposition 6.2.1*.

THEOREM 6.2.4. *Let f be Lipschitzian. The problem $\mathcal{P}(p,q)$ is calm at x_0 with modulus M' and radius ϵ if and only if the penalty function (6.2.3) is exact at x_0 on C for all $M \geq M'$.*

PROOF: Assume $\mathcal{P}(p,q)$ is calm at x_0 with modulus M' and radius ϵ. Given a point x near x_0 and in C, we define $\alpha := (g(x)-p)^+$ and $\beta := h(x)-q$. For the above defined $\alpha \in \mathbb{R}^s$ and $\beta \in \mathbb{R}^m$, x is feasible for $\mathcal{P}(p+\alpha, q+\beta)$. By *Proposition 6.2.1*, for any $M \geq M'$, we have $f(x_0) \leq f(x) + M|(\alpha,\beta)|$. Therefore

$$\begin{aligned} \phi_{M'}(x_0) &= f(x_0) + M'|((g(x_0)-p)^+, h(x_0)-q)| \\ &= f(x_0) \leq f(x) + M|(\alpha,\beta)| \qquad (6.2.4) \\ &= f(x) + M|((g(x)-p)^+, h(x)-q)|, \end{aligned}$$

i.e. the penalty function (6.2.3) is exact at x_0 on C for all $M \geq M'$.

On the other hand, if $\phi_M(x)$ is exact at x_0 for all $M \geq M'$, then for any given $\alpha \in \mathbb{R}^s$, $\beta \in \mathbb{R}^m$ and $x \in C \cap (x_0 + \epsilon\mathcal{B})$ for some $\epsilon > 0$ such that

$g(x) \leq p + \alpha$, and $h(x) = q + \beta$, we have

$$(g(x) - p)_i \leq |\alpha_i| \quad \text{for} \quad i = 1, \cdots, s;$$

$$|((g(x) - p)^+, h(x) - q)| \leq |(\alpha, \beta)|;$$

$$\begin{aligned} f(x_0) &\leq f(x) + M'|((g(x) - p)^+, h(x) - q)| \\ &\leq f(x) + M'|(\alpha, \beta)|. \end{aligned}$$

Hence $\mathcal{P}(p, q)$ is calm at x_0 with modulus M' and radius ϵ . □

For optimization problems that incorporate inequality constraints and an abstract constraint, we have a very nice property that *almost all problems are calm.* To be more precise, we have:

THEOREM 6.2.5. *([Clar83], 6.4.5) Consider $\mathcal{P}_1(p)$, i.e. problem (1.1.1) incorporating only inequality constraints $g(x) \leq p$ and an abstract constraint $x \in C$. Suppose that the value function $V(p)$ is finite for p near 0 . Then for almost all p in a neighborhood of 0 , the problem $\mathcal{P}_1(p)$ is calm.*

PROOF: By the definition of value function, in a neighborhood of 0 , the function $V(p) = V(p_1, \ldots, p_s)$ is finite and decreasing as a function of each component p_i . Such functions are differentiable almost everywhere with respect to *Lebesgue measure.* But $\mathcal{P}_1(p)$ is calm whenever V is differentiable at $p = (p_1, \ldots, p_s)$, so the result follows. □

It is natural to wonder if *Theorem 6.2.5* can be extended to the case where problem $\mathcal{P}$ incorporates equality constraints and inequality constraints with or without an abstract constraint (the problem is posed in [Clar83]). It is true that $\mathcal{P}(p, q)$ is calm for all (p, q) on a dense subset of any open set upon which V is bounded and lower semi-continuous.

THEOREM 6.2.6. *([Clar83], 6.4.6) If there is an open set* $\mathcal{U} \subset \mathbb{R}^s \times \mathbb{R}^m$ *such that* $V(p,q)$ *is bounded on* $\mathcal{U}$ *, then* $\mathcal{P}(p,q)$ *is calm on a dense subset of* $\mathcal{U}$ *.*

Unfortunately, a result similar to *Theorem 6.2.5* is not true for problems that incorporate equality constraints and an abstract constraint; counterexamples are provided in the next section.

6.3 Counterexamples

What makes equality constraints so different from inequality constraints? The proof of *Theorem 6.2.5* is easy because the value function is monotone and differentiable almost everywhere as the problem involves only inequality constraints and an abstract constraint. If we just look at the equality constraints and inequality constraints without considering their sensitivity, then equality constraints can be replaced by inequality constraints without loss of Lipschitz property and vice versa. For example, one equality constraint $h_i(x) = 0$ can be replaced by two inequality constraints: $h_i(x) \geq 0$, and $h_i(x) \leq 0$ or by a single inequality constraint $h_i^2(x) \leq 0$. But the first substitution results in the increasing of the number of constraints, and the sensitivity conducted in the two corresponding inequalities is not independent. The second substitution results in different meanings of sensitivity.

To provide a counterexample, we introduce a lemma about one dimensional nondifferentiable functions. Denote by

$$\overline{\varphi}^+(x) = \limsup_{y \to x^+} \frac{\varphi(y) - \varphi(x)}{y - x},$$

$$\underline{\varphi}^+(x) = \liminf_{y \to x^+} \frac{\varphi(y) - \varphi(x)}{y - x}$$

the right upper and the right lower Dini derivatives respectively.

LEMMA 6.3.1. *Let* $\varphi(x) : \mathbb{R} \to \mathbb{R}$ *be a continuous and nowhere differentiable function on a set* S *of positive measure. Then neither* $\overline{\varphi}^+(x)$ *nor* $\underline{\varphi}^+(x)$ *can be finite almost everywhere in* S .

PROOF: This is equivalent to *Theorem 4*, Chapter IX in [Saks33]. □

It is easy to find a nowhere differentiable continuous function $\varphi(x) : \mathbb{R} \to \mathbb{R}$. More specifically, there is a function whose directional derivative doesn't exist at any point in any direction. For example,

$$\varphi(x) = \sum_{n=0}^{\infty} (\sin 2^n x / 2^n)$$

is such a function. By the symmetric property of $\sin(x)$ and *Lemma 6.3.1*, neither $\overline{\varphi}^+(x)$ nor $\underline{\varphi}^+(x)$ can be finite almost everywhere in any positive measured set.

Example 6.3.1 Consider the set C in $\mathbb{R}^2$ given by

$$C = \{ (x, y) \mid y = \varphi(x), x \in (-\infty, \infty) \},$$

(where φ is defined above), and two functions $f(x, y), h(x, y) : \mathbb{R}^2 \to \mathbb{R}$ such that $f(x, y) = y$ and $h(x, y) = x$. It is trivial to see that C is a closed set in $\mathbb{R}^2$ and that f, h are all Lipschitz continuous. Then the following constrained minimization problem

$$\begin{cases} min\ f(x, y) \\ s.t.\ \ h(x, y) = p, \\ \qquad (x, y) \in C, \end{cases} \tag{6.3.1}$$

has the value function $V(p) = \varphi(p)$ and achieved at $(p, \varphi(p))$ for any $p \in \mathbb{R}$. So the value function $V(p)$ doesn't have finite right lower Dini derivative almost

everywhere in any neighborhood of 0 . i.e.

$$\begin{aligned} &\liminf_{\alpha \to 0} \frac{V(p+\alpha) - V(p)}{|\alpha|} \\ \leq &\liminf_{\alpha \to 0^+} \frac{V(p+\alpha) - V(p)}{|\alpha|} = \underline{\varphi}^+(x) \end{aligned} \tag{6.3.2}$$

can not be $> -\infty$ for p almost everywhere in any neighborhood of 0 . Therefore (6.3.1) cannot be calm for almost all p in any given neighborhood of 0 .

By examining the counterexample carefully, we know that its abstract constraint set is a closed set consisting of a nonsmooth curve. A variation on the above example also shows that the similar result is not true if the abstract constraint is removed and the Lipschitz condition on the constraint function is relaxed.

Example 6.3.2 Suppose φ and f are defined as above, and $h_1(x,y)$ is defined to be $h(x,y)$ in *Example 6.3.1*, then f, h_1 are Lipschitz functions. Define $h_2(x,y) = y - \varphi(x)$ which is from $\mathbb{R}^2$ to $\mathbb{R}$. Then the following constrained optimization problem

$$\begin{cases} min\ f(x,y) \\ s.t.\ \ h_1(x,y) = p_1, \\ \qquad h_2(x,y) = p_2, \end{cases} \tag{6.3.3}$$

has the value function $V(p_1,p_2) = p_2 + \varphi(p_1)$ and achieved at $(p_1,\ p_2 + \varphi(p_1))$ for any $(p_1,p_2) \in \mathbb{R}$. So the value function $V(p_1,p_2)$ doesn't have finite right

lower Dini derivatives almost everywhere in any neighborhood of 0, because

$$\begin{aligned}
&\liminf_{(\alpha_1,\alpha_2)\to 0} \frac{V(p_1+\alpha_1,p_2+\alpha_2)-V(p_1,p_2)}{|(\alpha_1,\alpha_2)|} \\
&= \liminf_{(\alpha_1,\alpha_2)\to 0} \frac{p_2+\alpha_2+\varphi(p_1+\alpha_1)-(p_2+\varphi(p_1))}{|(\alpha_1,\alpha_2)|} \\
&\le \liminf_{\substack{\alpha_1\to 0\\ \alpha_2=0}} \frac{p_2+\alpha_2+\varphi(p_1+\alpha_1)-(p_2+\varphi(p_1))}{|(\alpha_1,\alpha_2)|} \qquad (6.3.4)\\
&= \liminf_{\alpha_1\to 0} \frac{\varphi(p_1+\alpha_1)-\varphi(p_1)}{|\alpha_1|} \\
&\le \liminf_{\alpha_1\to 0^+} \frac{\varphi(p_1+\alpha_1)-\varphi(p_1)}{\alpha_1} = \underline{\varphi}^+(p_1).
\end{aligned}$$

which means that

$$\liminf_{(\alpha_1,\alpha_2)\to 0} \frac{V(p_1+\alpha_1,p_2+\alpha_2)-V(p_1,p_2)}{|(\alpha_1,\alpha_2)|} > -\infty \qquad (6.3.5)$$

can not hold for almost all (p_1,p_2) in any neighborhood of $(0,0)$ in $\mathbb{R}^2$.

It is not known whether an analogue of *Theorem 6.2.5* holds for Lipschitz optimization problems that incorporate only equality constraints (without an abstract constraint), i.e. for

$$\begin{cases} \min\ f(x) \\ s.t.\ \ h_j(x)=p_j, \quad \text{for} \quad j=1,\cdots,m. \end{cases}$$

The answer to the problem is very significant, at least from a theoretical point of view, in optimization theory and sensitivity analysis.

CHAPTER 7

TANGENT CONES AND NORMAL CONES TO THE CONSTRAINT SET

Before developing sufficient conditions for calmness, we take a look at the relationships between the normal cone of the constraint set and the generalized gradients of the constraint functions. These kinds of relationships have appeared in [Clar83], [RoWe93] for inequality constraints, and in [Rock85] for equality constraints. But all of them are inclusions, and equality holds only under very strong conditions (e.g. the constraint functions are assumed to be strictly differentiable or continuously differentiable). The results appearing in [Rock85] and [RoWe93] are obtained by using a *chain rule* developed in the literature. This chapter establishes stronger results about those relationships, which have important significance for themselves in both theory and applications.

7.1 Definitions and Review of the Results

We first give some definitions and review some of the results already available in the literature.

DEFINITION 7.1.1. *Let* $F : \mathbb{R}^n \to \mathbb{R}^m$. *We say* F *is differentiable at* x *if there exist a continuous linear operator from* $\mathbb{R}^n$ *to* $\mathbb{R}^m$, *denoted by* $JF(x)$, *such that*

$$\lim_{y \to x} \frac{F(y) - F(x) - JF(x)(y - x)}{|y - x|} = 0. \tag{7.1.1}$$

$JF(x)$ *is called the Jacobian of* F *at* x . *In the case* $m = 1$, *it is called the gradient and denoted by* $\nabla F(x)$.

We call a function $F : \mathbb{R}^n \to \mathbb{R}$ *strictly differentiable at* x *if for any* $v \in \mathbb{R}^n$ *, we have*

$$\lim_{\substack{y \to x \\ t \downarrow 0}} \frac{F(y+tv) - F(y)}{t} = \nabla F(x)v, \tag{7.1.2}$$

and the convergence is uniform for v *in compact sets (the "uniform" property is automatically true if* F *is Lipschitzian).*

From the definition, we know that the *strict differentiability* of F at x is a strong condition, which is implied by the continuous differentiability of F around x . The *strict differentiability* of a Lipschitz function F means that the *generalized Jacobian* of F at x , defined by (1.1.4), is a singleton and equals $\nabla F(x)$ (see 2.2.1 and 2.2.4 in [Clar83]).

For a sequence of subsets $\{C_i\}$ in $\mathbb{R}^n$, we define the upper and lower limits of $\{C_i\}$

$$\begin{aligned} \limsup_{i \to \infty} C_i &:= \{x \in \mathbb{R}^n : \liminf_{i \to \infty} dist(x, C_i) = 0\}, \\ \liminf_{i \to \infty} C_i &:= \{x \in \mathbb{R}^n : \limsup_{i \to \infty} dist(x, C_i) = 0\}. \end{aligned} \tag{7.1.3}$$

Obviously, $\liminf C_i \subset \limsup C_i$. If the above two sets are the same, then we say that $\lim C_i$ exists and is equal to the same set.

DEFINITION 7.1.2. *Let* C *be a subset in* $\mathbb{R}^n$ *and* $x \in C$ *. We call*

$$\begin{aligned} T_C(x) &:= \liminf_{\substack{y \to x, t \downarrow 0 \\ y \in C}} (C - y)/t, \\ K_C(x) &:= \limsup_{t \downarrow 0} (C - x)/t, \end{aligned}$$

the Clarke tangent cone of C *at* x *and the contingent cone of* C *at* x *respectively (see [RoWe93], §5A).*

DEFINITION 7.1.3. *A vector* v *is a proximal normal to* C *at a point* $x \in C$ *if for some* $\epsilon > 0$ *one has that* x *is a point of* C *nearest to the point* $x + \epsilon v$ *. We*

call the set of all proximal normals to C *at* $x \in C$ *the proximal normal cone of* C *at* x, *which is denoted by* $PN_C(x)$.

A vector $v \in \mathbb{R}^n$ *is a strict normal vector to a closed set* C *at* $x \in C$, *written* $v \in \widehat{N}_C(x)$, *if* $\langle v, y-x\rangle \le o|y-x|$ *for* $y \to x$ *and* $y \in C$, *i.e.*

$$\limsup_{\substack{y \to x,\\ y \neq x, y \in C}} \langle v, y-x\rangle / |y-x| \le 0.$$

The polar of $T_C(x)$:

$$N_C(x) := \{\, \xi \in \mathbb{R}^n : \langle \xi, v\rangle \le 0 \quad \textit{for all} \quad v \in T_C(x) \,\} \tag{7.1.4}$$

is called the Clarke normal cone of C *at* x.

There are quite a few equivalent definitions of the above concepts; these can be found in [Clar83] and [RoWe93]. It is well known that the Clarke tangent cone and the Clarke normal cone are closed convex cones and the contingent cone contains the Clarke tangent cone.

DEFINITION 7.1.4. *A set* C *is regular at* $x \in C$ *provided* $T_C(x) = K_C(x)$.

DEFINITION 7.1.5. *A function* $f : \mathbb{R}^n \to \mathbb{R}$ *is called regular at* $x \in C$ *provided*

(i) For all v, *the usual one-side directional derivative* $f'(x; v)$ *exists;*

(ii) For all v, $f'(x; v) = f^\circ(x; v)$.

THEOREM 7.1.6. *Let function* f *be Lipschitzian near* x,

(a) f *is strictly differentiable at* x *iff both* f *and* $-f$ *are regular at* x;

(b) If f *is convex, then* f *is regular at* x;

(c) A finite linear combination by nonnegative scalars of regular functions at x *is regular at* x.

PROOF: Results of (b) and (c) are from §2.3 in [Clar83]. The proof of (a) can

be found in [RoWe93], where a number of equivalent conditions of strict differentiability are established. The following is a direct proof of (a). We see first that strict differentiability does mean that f and $-f$ are both regular at x, and on the other hand, if f and $-f$ are both regular,

$$\begin{aligned}
&\liminf_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+tv)-f(y)}{t} = -\limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y)-f(y+tv)}{t}\\
&= -\limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y+tv-tv)-f(y+tv)}{t} = -f^\circ(x;-v)\\
&= -f'(x;-v) = -\lim_{t\downarrow 0} \frac{f(x-tv)-f(x)}{t}\\
&= \lim_{t\downarrow 0} \frac{(-f)(x-tv)-(-f)(x)}{t} = (-f)'(x;-v)\\
&= (-f)^\circ(x;-v) = \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{(-f)(y-tv)-(-f)(y)}{t}\\
&= \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y)-f(y-tv)}{t} = \limsup_{\substack{y\to x\\ t\downarrow 0}} \frac{f(y-tv+tv)-f(y-tv)}{t}\\
&= \limsup_{\substack{z\to x\\ t\downarrow 0}} \frac{f(z+tv)-f(z)}{t},
\end{aligned}$$

i.e. f is strictly differentiable at x as f is Lipschitzian. □

For a locally Lipschitz function $h : \mathbb{R}^n \to \mathbb{R}^m$ and a fixed $q \in \mathcal{R}(h)$, the range of $h(x)$, we denote

$$\begin{aligned}
C_L := C_L(q) &= \{\, x \in \mathbb{R}^n : h_i(x) \le q_i \quad \text{for } i = 1, \dots, m \,\},\\
C_U := C_U(q) &= \{\, x \in \mathbb{R}^n : h_i(x) \ge q_i \quad \text{for } i = 1, \dots, m \,\},\\
C := C(q) = C_L \cap C_U &= \{\, x \in \mathbb{R}^n : h_i(x) = q_i \quad \text{for } i = 1, \dots, m \,\}.
\end{aligned}$$

THEOREM 7.1.7. *([Clar83], 2.4.7) Let $m = 1$ and h be Lipschitzian near x, for $x \in C$, and suppose $0 \notin \partial h(x)$. Then one has*

$$\{\, v \in \mathbb{R}^n : h^\circ(x;v) \le 0 \,\} \subset T_{C_L}(x). \tag{7.1.5}$$

If h is regular at x, then equality holds and C_L is regular at x.

COROLLARY 7.1.8. *([Clar83], Corollary to 2.4.7) Under the conditions of Theorem 7.1.7, one has*

$$N_{C_L}(x) \subset \bigcup_{\lambda \geq 0} \lambda \partial h(x). \tag{7.1.6}$$

And if h is regular at x, then equality holds.

Now if we let $\varphi(x) = -h(x)$, and apply the previous theorem to $\varphi(x)$ (however the regularity of $\varphi(x)$ is generally not preserved for regular function $h(x)$), then from *Theorem 7.1.7*, with $C_U = \{ y \in \mathbb{R}^n : \varphi(y) \leq \varphi(x) \}$, we have

$$\{ v \in \mathbb{R}^n : \varphi^\circ(x; v) \leq 0 \} \subset T_{C_U}(x). \tag{7.1.7}$$

Because $\varphi^\circ(x; v) = (-h)^\circ(x; v) = h^\circ(x; -v)$, $\partial(-h)(x) = -\partial h(x)$,

$$\{ v \in \mathbb{R}^n : h^\circ(x; -v) \leq 0 \} = -\{ v \in \mathbb{R}^n : h^\circ(x; v) \leq 0 \},$$

then we have the following result:

COROLLARY 7.1.9. *Let $m = 1$ and h be Lipschitzian near x, and suppose $0 \notin \partial h(x)$. Then one has*

$$-\{ v \in \mathbb{R}^n : h^\circ(x; v) \leq 0 \} \subset T_{C_U}(x),$$

$$N_{C_U}(x) \subset \bigcup_{\lambda \leq 0} \lambda \partial h(x),$$

and equality holds if $-h$ is regular at x, and C_U is regular at x in this case.

The following result is from [Rock85], which will be used later in §7.3.

THEOREM 7.1.10. *Suppose $C = \cap_{i=1}^m C_i$ where C_i is closed in $\mathbb{R}^n$ for $i = 1, \ldots, m$. Let x be a point of C. Under the assumption that the only possible*

situation that $\sum_{i=1}^{m} v_i = 0$ *for* $v_i \in N_{C_i}(x)$ *is* $v_i = 0$ *for* $i = 1, \ldots, m$, *one has that*

$$T_C(x) \supset \bigcap_{i=1}^{m} T_{C_i}(x), \qquad N_C(x) \subset cl\Big(\sum_{i=1}^{m} N_{C_i}(x)\Big). \tag{7.1.8}$$

If in addition, the sets C_i *are all regular at* x , *then* C *is regular at* x *and the estimates hold as equalities.*

7.2 Results for One Equality Constraint

In this section, we consider the special case of one equality constraint. *Theorem 7.2.1* and *Corollary 7.2.3* have already appeared in [Hiri79]. A new proof for *Theorem 7.2.1* is provided here.

THEOREM 7.2.1. *Suppose* h *is continuous and* $m = 1$, *i.e. the case of one equality constraint. For* $x \in C$, *we have*

$$T_C(x) = T_{C_L}(x) \bigcap T_{C_U}(x). \tag{7.2.1}$$

PROOF: First we show that $T_C(x) \subset T_{C_L}(x) \cap T_{C_U}(x)$. By the symmetric property of C_L and C_U , it is sufficient to show that $T_C(x) \subset T_{C_L}(x)$. By definition of the Clarke tangent cone, $T_C(x)$ is given by

$$\liminf_{\substack{y \to x, t \downarrow 0 \\ y \in C}} (C - y)/t = \{\, v : \limsup_{\substack{y \to x, t \downarrow 0 \\ y \in C}} dist\,(v, (C - y)/t) = 0 \,\}. \tag{7.2.2}$$

It is easy to see that

$$dist\,(v, (C - y)/t) = \frac{1}{t} dist\,(C, y + tv)\,.$$

Then $v \in T_C(x)$, iff

$$\limsup_{\substack{y \to x, t \downarrow 0 \\ y \in C}} \frac{1}{t} dist\,(C, y + tv) = 0. \tag{7.2.3}$$

Now consider $y_i \subset C_L$, $y_i \to x$ and $t_i \downarrow 0$. Let $S_i := \{ y_i + sv : 0 \le s \le t_i \}$, i.e. a closed interval, then either $y_i + t_i v \in C_L$ or $y_i + t_i v \notin C_L$. In the first case,

$$dist\,(v, (C_L - y_i)/t_i) = \frac{1}{t_i} dist\,(C_L, y_i + t_i v) = 0, \tag{7.2.4}$$

In the second case, as $m = 1$, there exists $y_i' \in S_i \cap C$ such that $y_i + t_i v = y_i' + \alpha_i v$ for some $\alpha_i \in [0, t_i]$, and $dist(C_L, y_i + t_i v) = dist(C, y_i' + \alpha_i v)$. Therefore

$$\begin{aligned} & dist\,(v, (C_L - y_i)/t_i) = \frac{1}{t_i} dist\,(C_L, y_i + t_i v) \\ =& \frac{1}{t_i} dist\,(C, y_i' + \alpha_i v) = \frac{1}{\alpha_i} dist\,(C, y_i' + \alpha_i v) \cdot \frac{\alpha_i}{t_i}, \\ \le& \frac{1}{\alpha_i} dist\,(C, y_i' + \alpha_i v). \end{aligned} \tag{7.2.5}$$

As $y_i' \in [y_i, y_i + t_i v]$, then $y_i \in C_L$, $y_i \to x$ and $t_i \downarrow 0$ imply $y_i' \to x$ and $\alpha_i \downarrow 0$.

Combining the two cases, we have either

$$\limsup_{i \to \infty} dist\,(v, (C_L - y_i)/t_i) = 0,$$

which happens if $y_i + t_i v$ is eventually in C_L , or

$$\begin{aligned} & \limsup_{i \to \infty} dist\,(v, (C_L - y_i)/t_i) \\ \le & \limsup_{i \to \infty} \frac{1}{\alpha_i} dist\,(C, y_i + \alpha_i v) \\ \le & \limsup_{\substack{y' \to x, \alpha \downarrow 0 \\ y' \in C}} \frac{1}{\alpha} dist\,(C, y' + \alpha v) = 0, \end{aligned}$$

as $v \in T_C(x)$ and $0 \le \alpha \le t_i$, which proves that $v \in T_{C_L}(x)$.

On the other hand, by (7.2.2), if $v \in T_{C_L}(x) \cap T_{C_U}(x)$ and as $C = C_L \cap C_U$, then for any $y \in C$ and $t > 0$, there exist $u'_{y,t} \in C_L$ and $u''_{y,t} \in C_U$ such that $\left|(u'_{y,t} - y)/t - v\right| \to 0$ and $\left|(u''_{y,t} - y)/t - v\right| \to 0$ as $y \in C, y \to x, t \downarrow 0$. By the definition of C, C_L and C_U and the continuity of $h(x)$, $\exists\, \lambda_{y,t} \in [0,1]$ s.t.

$u_{y,t} = \lambda_{y,t} \cdot u'_{y,t} + (1 - \lambda_{y,t}) \cdot u''_{y,t} \in C$, and obviously $|(u_{y,t} - y)/t - v| \to 0$ as $y \in C$, $y \to x$, $t \downarrow 0$, i.e. $v \in T_C(x)$, which proves the theorem. □

In the same way, we can prove a similar result for contingent cones instead of Clarke tangent cones:

THEOREM 7.2.2. *Suppose* h *is continuous and* $m = 1$. *For* $x \in C$, *we have*

$$K_C(x) = K_{C_L}(x) \bigcap K_{C_U}(x). \tag{7.2.6}$$

Remark: *Theorem 7.2.1* is equivalent to saying that for any two nonempty closed sets C_1 and C_2 in $\mathbb{R}^n$, if $C_1 \cup C_2 = \mathbb{R}^n$, then for any $x \in C := C_1 \cap C_2$, we have (7.2.1). For the constraint functions h , all we require is that $h(x)$ be continuous. We need neither a regularity condition nor a Lipschitz condition on the constraint function. This means that the theorem may be very useful for other applications. For example, the theorem can be used to provide necessary optimality conditions for a mathematical program with one equality constraint; see [MiPe92].

Immediately from *Theorem 7.2.1*, we have,

COROLLARY 7.2.3. *If* $m = 1$ *and* $h : \mathbb{R}^n \to \mathbb{R}$ *is continuous, then*

$$N_C(x) = cl(N_{C_L}(x) + N_{C_U}(x)). \tag{7.2.7}$$

Now combining the above results, we have

THEOREM 7.2.4. *Suppose* $m = 1$ *and* h *is Lipschitz around* x . *If* $0 \notin \partial h(x)$, *then*

$$\begin{aligned} T_C(x) &\supset \{ v \in \mathbb{R}^n : h^\circ(x; v) \le 0; (-h)^\circ(x; v) \le 0 \} \\ &= \{ v \in \mathbb{R}^n : \langle v, \xi \rangle = 0 \quad \text{for all} \quad \xi \in \partial h(x) \}, \end{aligned} \tag{7.2.8}$$

$$N_C(x) \subset span\{\partial h(x)\}, \tag{7.2.9}$$

and all the inclusions can be changed to equalities if h *is strictly differentiable at* x *. In the latter case,* C *is regular at* x *and* $N_C(x) = span\{\nabla h(x)\}$ *. Here* $span\{\partial h(x)\}$ *means the subspace generated by* $\partial h(x)$ *.*

PROOF: The inclusion of (7.2.8) follows from *Theorem 7.2.1*, *Theorem 7.1.7* and *Corollary 7.1.9*. The equality of (7.2.8) holds because v satisfies $h^\circ(x,v) \leq 0$ and $(-h)^\circ(x,v) \leq 0$ iff $\langle v, \xi \rangle = 0$ for all $\xi \in \partial h(x)$. The inclusion in (7.2.9) is true because $N_C(x)$ is the polar of $T_C(x)$. The result for the special case holds due to the special cases in *Theorem 7.1.7* and *Corollary 7.1.9*, and the regularity of C follows by

$$T_C(x) = T_{C_L}(x) \bigcap T_{C_U}(x) = K_{C_L}(x) \bigcap K_{C_U}(x) = K_C(x).$$

□

Remark: (7.2.9) has already appeared in [Rock85].

We end this section by presenting a result about the contingent cone to the set of one equality constraint at a differentiable point.

THEOREM 7.2.5. *Suppose* $m = 1$ *and* h *is continuous. If* h *is differentiable at* x *with* $x \in C$ *, and* $\nabla h(x) \neq 0$ *, then*

$$K_C(x) = \{\, v \in \mathbb{R}^n : v^T \nabla h(x) = 0 \,\}. \tag{7.2.10}$$

PROOF: Let $v \in K_C(x)$, and $|v| = 1$ then there exists $\{y_i\} \subset C$ such that $y_i \to x$ and $(y_i - x)/|y_i - x| \to v$. Because h is differentiable at x , then

$$0 = h(y_i) - h(x) = \nabla h(x)^T (y_i - x) + o|y_i - x|,$$

dividing the above expression by $|y_i - x|$ and taking limit as $i \to \infty$, we have $\nabla h(x)^T v = 0$.

On the other hand, let $\nabla h(x)^T v = 0$ for some v. For any positive integer i, we consider $y' := x + t\big(v + \frac{1}{i}\nabla h(x)\big)$ and $y'' := x + t\big(v - \frac{1}{i}\nabla h(x)\big)$. As h is differentiable at x, then

$$h(y') - h(x) = \frac{1}{i}t|\nabla h(x)|^2 + o(t|v + \frac{1}{i}\nabla h(x)|),$$

which is greater than 0 for t small enough. And

$$h(y'') - h(x) = -\frac{1}{i}t|\nabla h(x)|^2 + o(t|v + \frac{1}{i}\nabla h(x)|),$$

which is less than 0 for t small enough. Therefore there exists $t_i > 0$ such that $h(y_i') - h(x) > 0$ and $h(y_i'') - h(x) < 0$ for $y_i' := x + t_i\big(v + \frac{1}{i}\nabla h(x)\big)$ and $y_i'' := x + t_i\big(v - \frac{1}{i}\nabla h(x)\big)$. By the *intermediate value theorem*, along the line segment in between y_i' and y_i'', there exists y_i such that $h(y_i) - h(x) = 0$, i.e. $y_i \in C$. Obviously, we have

$$\frac{y_i - x}{|y_i - x|} \to \frac{v}{|v|},$$

which means $v \in K_C(x)$ and the theorem is proved. □

Remark: For inequality constraint set $C_U = \{y : h(y) \geq h(x)\}$, we have the similar result. If h is a continuous function and differentiable at x with $\nabla h(x) \neq 0$, then

$$K_{C_U}(x) = \{\, v \in \mathbb{R}^n : v^T \nabla h(x) \geq 0 \,\}. \tag{7.2.11}$$

7.3 The Clarke Normal Cone to the Equality Constraint Set

Theorem 7.1.7 and *Corollary 7.1.9* can be easily generalized to the case of an arbitrary integer m. The proofs are very similar to those provided by Clarke for the case of $m = 1$. Instead of assuming $0 \notin \partial h(x)$, which ensures that there exists v s.t. $h^\circ(x;v) < 0$ in *Theorem 7.1.7*, we assume that there exists $v_0 \in \mathbb{R}^n$ s.t. $h_i^\circ(x;v_0) < 0$ for $i = 1, \ldots, m$ in *Theorem 7.3.1*. We denote

$$I = \{ \lambda \in \mathbb{R}^m : \lambda_i \geq 0 \quad \text{for} \quad i = 1, \cdots, m, \quad \sum_{i=1}^{m} \lambda_i = 1 \}, \tag{7.3.1}$$

and

$$h_\lambda(x) = \sum_{i=1}^{m} \lambda_i h_i(x) \quad \text{for} \quad \lambda \in \mathbb{R}^m. \tag{7.3.2}$$

THEOREM 7.3.1. *Consider an arbitrary integer m. Let $h(x)$ be Lipschitzian near x, and suppose there exists $v_0 \in \mathbb{R}^n$ s.t. $h_i^\circ(x;v_0) < 0$ for $i = 1, \cdots, m$. Then one has*

$$\{ v \in \mathbb{R}^n : h_i^\circ(x;v) \leq 0 \quad \text{for} \quad i = 1, \ldots, m \} \subset T_{C_L}(x), \tag{7.3.3}$$

which is equivalent to

$$\{ v \in \mathbb{R}^n : h_\lambda^\circ(x;v) \leq 0 \quad \text{for} \quad \lambda \in I \} \subset T_{C_L}(x). \tag{7.3.4}$$

And if h_i is regular at x for $i = 1, \ldots, m$, then equalities hold for (7.3.3) and (7.3.4), and C_L is regular at x.

PROOF: The proof of the theorem except (7.3.4) is similar to the one provided by Clarke for the case of $m = 1$ (see 2.4.7 in [Clar83]). (7.3.4) is equivalent to (7.3.3), because the left hand side of (7.3.4) is obviously included in the left hand side of (7.3.3). On the other hand, if v belongs to the left hand side of (7.3.3), then by the definition of the generalized gradient,

$$h_\lambda^\circ(x;v) \leq \sum_{i=1}^{m} \lambda_i h_i^\circ(x;v) \leq 0$$

for all $\lambda \in I$, i.e. v also belongs to the left hand side of (7.3.4). □

COROLLARY 7.3.2. *Under the condition of Theorem 7.3.1, one has*

$$N_{C_L}(x) \subset \bigcup_{\alpha \geq 0} \alpha \partial h_1(x) + \cdots + \bigcup_{\alpha \geq 0} \alpha \partial h_m(x), \tag{7.3.5}$$

which is equivalent to

$$N_{C_L}(x) \subset cl\Big(co\big(\bigcup_{\lambda \in I} \bigcup_{\alpha \geq 0} \alpha \partial h_\lambda(x)\big)\Big). \tag{7.3.6}$$

And if h_i is regular at x for $i = 1, \ldots, m$, then equality holds.

PROOF: First, we notice that the left hand sides of (7.3.3) and (7.3.4) are equivalent to

$$\bigcap_{i=1}^{m} \{\, v \in \mathbb{R}^n : h_i^\circ(x; v) \leq 0 \,\}$$

and

$$\bigcap_{\lambda \in I} \{\, v \in \mathbb{R}^n : h_\lambda^\circ(x; v) \leq 0 \,\}$$

respectively. Second, for $i = 1, \cdots, m$, $\lambda \in I$, $\{\, v \in \mathbb{R}^n : h_i^\circ(x; v) \leq 0 \,\}$, $\{\, v \in \mathbb{R}^n : h_\lambda^\circ(x; v) \leq 0 \,\}$ are all closed sets because h is Lipschitzian. Third, the polar cones of $\{\, v \in \mathbb{R}^n : h_i^\circ(x; v) \leq 0 \,\}$ and $\{\, v \in \mathbb{R}^n : h_\lambda^\circ(x; v) \leq 0 \,\}$ are precisely the closed convex cones generated by $\partial h_i(x)$ and $\partial h_\lambda(x)$, i.e., $\cup_{\alpha \geq 0} \alpha \partial h_i(x)$ and $\cup_{\alpha \geq 0} \alpha \partial h_\lambda(x)$ respectively (which are already closed). Fourth, $\cup_{i=1}^m \partial h_i(x)$ doesn't contain zero. Otherwise there exists $v_i \in \partial h_i(x)$, for $i = 1, 2, \cdots, m$, such that $v_1 + v_2 + \cdots + v_m = 0$, i.e., $v_1 = -\sum_{i=2}^m v_i$. Because

$$0 > h_i^\circ(x; v_0) = \sup\{\, \langle v_0, \xi \rangle : \xi \in \partial h_i(x) \,\} \geq \langle v_0, v_i \rangle, \quad \text{for} \quad i = 1, 2, \cdots, m,$$

then

$$\langle v_0, v_1 \rangle = \langle v_0, -\sum_{i=2}^{m} v_i \rangle = -\sum_{i=2}^{m} \langle v_0, v_i \rangle > 0,$$

which is a contradiction.

Therefore (7.3.5) follows from (7.3.3) by *Corollary 16.4.2* in [Rock70] and the fact that the right hand side of (7.3.5) is already closed because $\cup_{i=1}^{m} \partial h_i(x)$ is closed and does not contain zero. And (7.3.6) follows from (7.3.4) by *Corollary 16.5.2* in [Rock70]. □

Remark 1: The above two results are actually true for Banach space instead of finite dimensional spaces. And *Corollary 7.3.2* is stronger than *Corollary 2* of *Theorem 2.4.7* in [Clar83] because if h_i is strictly differentiable at x for $i = 1, \ldots, m$ and $\nabla h_1(x), \ldots, \nabla h_m(x)$ are linearly independent, then h_i is regular at x and there exists $v_0 \in \mathbb{R}^n$ s.t. $\nabla h_i(x)^T v_0 < 0$ for $i = 1, \ldots, m$, i.e. $h_i^\circ(x; v_0) < 0$ for $i = 1, \ldots, m$, and (7.3.5) and (7.3.6) hold for equality.

The above results can not be generalized to the situation of an equality constraint set by simply converting an equality to an inequality, such as

$$\{\, y : h(y) = h(x) \,\} = \{\, y : (h(y) - h(x))^2 \leq 0 \,\},$$

because the condition corresponding to $0 \notin \partial h(x)$ in *Theorem 7.1.7* is not satisfied in this case. To solve this problem, one alternative is to use the results in §7.2 and *Theorem 7.1.10*.

THEOREM 7.3.3. *Let* $h = (h_1, \ldots, h_m) : \mathbb{R}^n \to \mathbb{R}^m$, *and* $x \in C$. *Suppose* h *is Lipschitzian around* x, *and there are no* $v_i \in \partial h_i(x)$ *for* $i = 1, \ldots, m$ *s.t.* $v_1, \ldots v_m$ *are linearly dependent, then*

$$\begin{aligned}
&T_C(x) \\
\supset &\{\, v \in \mathbb{R}^n : h_i^\circ(x; v) \leq 0;\ (-h_i)^\circ(x; v) \leq 0 \quad \textit{for all} \quad i = 1, \cdots, m \,\} && (7.3.7) \\
= &\{\, v \in \mathbb{R}^n : \langle v, \xi \rangle = 0 \quad \textit{for all} \quad \xi \in \cup_{i=1}^{m} \partial h_i(x) \,\} && (7.3.8) \\
= &\{\, v \in \mathbb{R}^n : h_\lambda^\circ(x; v) \leq 0 \quad \textit{for all} \quad \lambda \in \mathcal{B}^m \,\}, && (7.3.9)
\end{aligned}$$

$$N_C(x) \subset cl\Big(co\big(\bigcup_{\lambda\in\mathcal{B}^m}\bigcup_{\alpha\geq 0}\alpha\partial h_\lambda(x)\big)\Big) \tag{7.3.10}$$

$$= \{\sum_{i=1}^{m}\lambda_i\partial h_i(x) \,:\, \lambda_i\in\mathbb{R}, i=1,\ldots,m\} \tag{7.3.11}$$

$$=: span\{\partial h_i(x), i=1,\ldots,m\}.$$

All the above inclusions can be changed to equalities if h is strictly differentiable at x. In this case,

$$N_C(x) = \{\sum_{i=1}^{m}\lambda_i\nabla h_i(x) : (\lambda_1,\ldots,\lambda_m)\in\mathbb{R}^m\}.$$

PROOF: Denote $C^i = \{y \,:\, h_i(y) = h_i(x)\}$. Then C^i is closed. By the assumption, we have $0\notin\partial h_i(x)$, and that the conditions of *Theorem 7.2.4* are satisfied for C^i. (7.3.7) and the result for the special case that h is strictly differentiable at x follow from *Theorem 7.1.10* as $C = \cap_{i=1}^m C^i$.

The equivalence between (7.3.7) and (7.3.8) is the result of the equivalence between

$$h_i^\circ(x;v)\leq 0; \quad (-h_i)^\circ(x;v)\leq 0$$

and

$$\langle v,\xi\rangle = 0 \quad \text{for all} \quad \xi\in\partial h_i(x).$$

And the equivalence between (7.3.8) and (7.3.9) can be proved as in *Theorem 7.3.1*, and (7.3.10) and (7.3.11) can be proved as in *Corollary 7.3.2* (notice that a linear subspace in finite dimensional space is closed). □

Remark 2: (7.3.10) in *Theorem 7.3.3* has already appeared in [Rock85] as a result of a chain rule for the generalized gradient under the same condition.

Remark 3: Combining *Theorem 7.3.1*, *Corollary 7.3.2* and *Theorem 7.3.3*, we can obtain a similar result for a set determined by equality constraints and inequality constraints.

7.4 The Contingent Cone and the Strict Normal Cone at a Differentiable Point

In this section we present a few results about the contingent cone and the strict normal cone to the set of equality constraints at a differentiable point. We also establish necessary and sufficient conditions for proximal normality in terms of the Jacobian of the constraint functions.

THEOREM 7.4.1. *Suppose that* $h : \mathbb{R}^n \to \mathbb{R}^m$ *is a continuous function and is differentiable at* x . *If the Jacobian* $Jh(x)$ *is of full row rank (i.e. with rank* m *), then*

$$K_C(x) = \bigcap_{i=1}^{m} K_{C^i}(x) = \{ v \in \mathbb{R}^n \ : \ Jh(x)v = 0 \}, \tag{7.4.1}$$

where $C^i = \{ y \ : \ h_i(y) = h_i(x) \}$ *for* $i = 1, 2, \cdots, m$, *and* $C = \cap C^i = \{ y \ : \ h_i(y) = h_i(x), \ \text{for } i = 1, \cdots, m \}$ *which are all closed.*

PROOF: We prove the result for the case of $m = 2$, because the general case can be proved similarly. We only have to prove $K_C(x) \supset K_{C^1}(x) \cap K_{C^2}(x)$, as the other direction of the inclusion is trivial, and the second equality in (7.4.1) is a direct consequence of the first equality and *Theorem 7.2.5*. Let $v \in K_{C^1}(x) \cap K_{C^2}(x)$. Then from *Theorem 7.2.5*, $v^T \nabla h_1(x) = 0$, $v^T \nabla h_2(x) = 0$. By the assumption, $\nabla h_1(x)$ and $\nabla h_2(x)$ are linearly independent, so $|\langle \nabla h_1(x), \nabla h_2(x) \rangle|^2 < |\nabla h_1(x)|^2 |\nabla h_2(x)|^2$; i.e. there exists β_0 such that

$$\frac{|\langle \nabla h_1(x), \nabla h_2(x) \rangle|}{|\nabla h_2(x)|^2} < \beta_0 < \frac{|\nabla h_1(x)|^2}{|\langle \nabla h_1(x), \nabla h_2(x) \rangle|}.$$

Choosing $\alpha_0 = 1$, then for $u(\alpha,\beta) := \alpha\nabla h_1(x) + \beta\nabla h_2(x)$, we have

$$\begin{aligned}
&\nabla h_1(x)^T u(\ \alpha_0,\beta) > 0 \text{ for all } \beta \in (-\beta_0,\beta_0);\\
&\nabla h_1(x)^T u(-\alpha_0,\beta) < 0 \text{ for all } \beta \in (-\beta_0,\beta_0);\\
&\nabla h_2(x)^T u(\ \alpha,\ \beta_0) > 0 \text{ for all } \alpha \in (-\alpha_0,\alpha_0);\\
&\nabla h_2(x)^T u(\alpha,-\beta_0) < 0 \text{ for all } \alpha \in (-\alpha_0,\alpha_0).
\end{aligned} \tag{7.4.2}$$

For any positive integer i , denote $y_{t,\alpha,\beta} := x + t\big(v + \frac{1}{i}u(\alpha,\beta)\big)$. Then we have

$$\begin{aligned}
h_1(y_{t,\alpha_0,\beta}) - h_1(x) &= \frac{1}{i}t\nabla h_1(x)^T u(\alpha_0,\beta) + o\big(t|v + \frac{1}{i}u(\alpha_0,\beta)|\big);\\
h_1(y_{t,-\alpha_0,\beta}) - h_1(x) &= \frac{1}{i}t\nabla h_1(x)^T u(-\alpha_0,\beta) + o\big(t|v + \frac{1}{i}u(-\alpha_0,\beta)|\big);\\
h_2(y_{t,\alpha,\beta_0}) - h_2(x) &= \frac{1}{i}t\nabla h_2(x)^T u(\alpha,\beta_0) + o\big(t|v + \frac{1}{i}u(\alpha,\beta_0)|\big);\\
h_2(y_{t,\alpha,-\beta_0}) - h_2(x) &= \frac{1}{i}t\nabla h_2(x)^T u(\alpha,-\beta_0) + o\big(t|v + \frac{1}{i}u(\alpha,-\beta_0)|\big).
\end{aligned} \tag{7.4.3}$$

By the definition of differentiability, each little " o " is uniformly small relative to t with respect to $\alpha \in [-\alpha_0,\alpha_0]$ and $\beta \in [-\beta_0,\beta_0]$. So by (7.4.2) there exist t_i such that

$$\begin{aligned}
&h_1(y_{t_i,\alpha_0,\beta}) - h_1(x) > 0 \text{ for all } \beta \in (-\beta_0,\beta_0);\\
&h_1(y_{t_i,-\alpha_0,\beta}) - h_1(x) < 0 \text{ for all } \beta \in (-\beta_0,\beta_0);\\
&h_2(y_{t_i,\alpha,\beta_0}) - h_2(x) > 0 \text{ for all } \alpha \in (-\alpha_0,\alpha_0);\\
&h_2(y_{t_i,\alpha,-\beta_0}) - h_2(x) < 0 \text{ for all } \alpha \in (-\alpha_0,\alpha_0).
\end{aligned} \tag{7.4.4}$$

Therefore by the *intermediate value theorem* (See *Lemma 4* in [MuWi74], the general case can be proved very easily by using topological degree (Brouwer degree); the proof is presented in the **Appendix** of this thesis), there exist $\alpha' \in [-\alpha_0,\alpha_0]$ and $\beta' \in [-\beta_0,\beta_0]$, such that $h_1(y_{t_i,\alpha',\beta'}) = h_1(x)$ and $h_2(y_{t_i,\alpha',\beta'}) = h_2(x)$, which means $y_i := y_{t_i,\alpha',\beta'} \in C$. It is easy to check that

$$\frac{y_i - x}{|y_i - x|} \to \frac{v}{|v|}. \tag{7.4.5}$$

i.e. $v \in K_C(x)$, which proves the theorem. □

Notice that the first equality of the theorem is usually not true for arbitrary sets C^1. Examples can be easily provided. This result improves the result of *Theorem 5A.22* in [RoWe93], where h_i are supposed to be of the class $\mathcal{C}^1$.

With a similar proof, it is easy to generalize the above theorem to a constraint set determined by both equality and inequality constraints. Denote

$$D := \{\, y : h_i(y) = p_i \quad \text{for} \quad i = 1, \cdots, m;$$
$$g_j(y) \geq q_j \quad \text{for} \quad j = 1, \cdots, s \,\}.$$

THEOREM 7.4.2. *Suppose that* $h_i : \mathbb{R}^n \to \mathbb{R}$ *and* $g_j : \mathbb{R}^n \to \mathbb{R}$ *are continuous functions and differentiable at* x *for* $i = 1, \cdots, m$ *and* $j = 1, \cdots, s$. *If the* $\nabla h_1(x), \cdots, \nabla h_m(x), \nabla g_1(x), \cdots, \nabla g_s(x)$ *are linearly independent, then we have*

$$K_D(x) = \{\, v : \nabla h_i(x)^T v = 0 \quad \text{for} \quad i = 1, \cdots, m;$$
$$\nabla g_j(x)^T v \geq 0 \quad \text{for} \quad j = 1, \cdots, s \text{ with } g_j(x) = q_j \,\}.$$

As we know the strict normal cone is the polar of the contingent cone. Therefore, by using *Theorem 7.2.5*, we have:

THEOREM 7.4.3. *Under the condition of Theorem 7.4.2, we have*

$$\begin{aligned} \widehat{N}_C(x) = \{ & \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{j=1}^{s} \mu_j \nabla g_j(x) : \\ & \lambda_i \in \mathbb{R}, \mu_j \leq 0 \text{ with } \mu_j(g_j(x) - q_j) = 0 \,\}, \end{aligned} \tag{7.4.6}$$

where $\widehat{N}_C(x)$ *is the strict normal cone of* C *at* x *defined in Definition 7.1.3.*

The similar results for smooth constraint functions—to be more specific, the similar expressions for the Clarke tangent (normal) cone of a smooth manifold—have appeared in [AuFr90] and [RoWe93]. The similar results under the condition of *strict differentiability*, which is defined in (7.1.2), have also appeared in [Clar83].

Recall that a proximal normal is also a strict normal. Therefore we have:

COROLLARY 7.4.4. *Suppose that* $h : \mathbb{R}^n \to \mathbb{R}^m$ *is a continuous function and is differentiable at* x . *If* v *is a proximal normal of* C *at* x , *then* $v = Jh(x)^T\lambda$ *for some* $\lambda \in \mathbb{R}^m$.

Corollary 7.4.4 establishes a necessary condition for a proximal normal in terms of the Jacobian of the constraint functions at a differentiable point. With the help of the notion of the *proximal subgradient*, defined by Rockafellar in [Rock81], we can establish a sufficient condition for a proximal normal in terms of the Jacobian of the constraint functions.

DEFINITION 7.4.5. *For* $f : \mathbb{R}^n \to \mathbb{R}^m$, *a vector* u *in* $\mathbb{R}^n$ *is a proximal subgradient to* f *at* x *if for some positive* α ,

$$f(y) \geq f(x) + \langle u, y - x\rangle - \alpha|y - x|^2 \tag{7.4.7}$$

holds for y *in a neighborhood of* x . *That is*

$$\liminf_{y \to x} \frac{f(y) - f(x) - \langle u, y - x\rangle}{|y - x|^2} > -\infty. \tag{7.4.8}$$

Recall that h_λ is defined in (7.3.2), i.e. $h_\lambda = \sum_{i=1}^m \lambda_i h_i$ for some $\lambda \in \mathbb{R}^m$.

THEOREM 7.4.6. *Suppose that* h *is a continuous function and is differentiable at* x . *If* $(-h_\lambda)$ *has a proximal subgradient* $-u$ *at* x , *or equivalently*

$$\limsup_{y \to x} \frac{h_\lambda(y) - h_\lambda(x) - \langle u, y - x\rangle}{|y - x|^2} < \infty, \tag{7.4.9}$$

then $u = Jh(x)^T\lambda$ *and is also a proximal normal to* C *at* x .

PROOF: By *Proposition 1.3* in [Poli88], $u = \nabla h_\lambda(x) = Jh(x)^T\lambda$. From the condition (7.4.9),

$$\limsup_{\substack{y \to x \\ y \in C}} \frac{\langle u, y - x\rangle}{|y - x|^2} < \infty,$$

which is equivalent to saying that u is also a proximal normal to C at x. $\square$

Remark 1: The above two results are significant to characterize the proximal normals for equality constraint sets.

Remark 2: Generalization of the results in this section to usual Banach space is challenging as we used the *intermediate value theorem* to prove *Theorem 7.4.1* in finite dimensional space. Instead of "*differential*" in $\mathbb{R}^n$, we should use "*Hadamard differential*" in Banach space.

7.5 More Results about the Clarke Normal Cone

In §3, an inclusion expression for the Clarke normal cone was given. By using *Theorem 7.4.3*, we can prove equality instead of inclusion under one more condition.

Recall that if $h(x) : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous, then by Rademacher's Theorem, h is differentiable almost everywhere. Denote by Ω_h the set of points where h is not differentiable. Then from [Clar83], the generalized Jacobian $\partial h(x)$ of h at x is defined by

$$\partial h(x) = co\{\mathcal{J}(x)\} := co\{ \lim_{i \to \infty} Jh(x_i) : x_i \to x, x_i \notin \Omega_h \}. \tag{7.5.1}$$

The set $\mathcal{J}(x)$ is called the limiting Jacobian of h at x It is possible that $\partial h(x)$ contains non-full-rank matrices while $\mathcal{J}(x)$ does not, because $\partial h(x)$ is the closed convex hull of $\mathcal{J}(x)$. A typical example is: $h : \mathbb{R} \to \mathbb{R}$, $h(x) = |x|$, at $x = 0$.

Before proving our final result in this chapter, we need the following lemma:

LEMMA 7.5.1. *Let h be Lipschitzian around x. Then for any $\lambda \in \mathbb{R}^m$, h_λ, defined by (7.3.2), is Lipschitzian around x and*

$$\partial h_\lambda(x) = co\{\, J^T\lambda : J \in \mathcal{J}(x)\,\}, \tag{7.5.2}$$

where $\mathcal{J}(x)$ is the set of limiting Jacobian defined in (7.5.1).

PROOF: Take $h_\lambda(x)$ as a composition of two functions: $h_\lambda(x) = \phi \circ h(x)$ with $\phi : \mathbb{R}^m \to \mathbb{R}$ such that $\phi(y) = \lambda_1 y_1 + \cdots + \lambda_m y_m$. Because ϕ is strictly differentiable everywhere, the special case of 2.6.6 in [Clar83] applies, and the proof is finished. □

LEMMA 7.5.2. *Let h be Lipschitzian around x. If there is no $v_i \in \partial h_i(x)$ for $i = 1, \cdots, m$ such that $v_1, \cdots, v_m$ are linearly dependent, then $\mathcal{J}(x)$ contains only full-rank matrices.*

PROOF: Let $J \in \mathcal{J}(x)$. Then J is an $m \times n$ matrix and the i-th row is an element of $\partial h_i(x)$ by (1.1.4). Therefore, by the condition of the Lemma, J is of full rank. □

Before proving the main result of this section, let's define a new notion of the limiting gradient (Jacobian). Let S be a subset of $\mathbb{R}^n$. The *S-relative limiting gradient (Jacobian)* of h at x, denoted by $\mathcal{J}_S(x)$, consists of the cluster points of $Jh(y)$ for y in S approaching x. That is

$$\mathcal{J}_S(x) := \{\, \lim Jh(y_i) : y_i \in S,\ y_i \to x\,\}.$$

Immediately from the definition, we know $\mathcal{J}_S(x) = \mathcal{J}(x)$, if x lies in the interior of S; and $\mathcal{J}_S(x) = \emptyset$, if x lies outside of the closure of S.

THEOREM 7.5.3. *Let h be Lipschitzian around x. Suppose there is no $v_i \in$*

$\partial h_i(x)$ for $i = 1, \cdots, m$ such that $v_1, \cdots, v_m$ are linearly dependent, and that

$$\mathcal{J}_C(x) = \mathcal{J}(x). \tag{7.5.4}$$

Then we have

$$\begin{aligned} N_C(x) &= co\{\, J^T\lambda : \lambda \in \mathbb{R}^m \text{ and } J \in \partial h(x)\,\} \\ &= co\{\, J^T\lambda : \lambda \in \mathbb{R}^m \text{ and } J \in \mathcal{J}(x)\,\}. \end{aligned} \tag{7.5.5}$$

PROOF: By using *Theorem 7.3.3* and *Lemma 7.5.1*, we have

$$\begin{aligned} & N_C(x) \\ \subset & cl\Big(co\Big(\bigcup_{\lambda\in\mathcal{B}^m}\bigcup_{\alpha\geq 0}\alpha\partial h_\lambda(x)\Big)\Big) \\ = & cl\big(co\{\,\alpha J^T\lambda : \alpha \geq 0, J \in \partial h(x), \lambda \in \mathcal{B}^m\,\}\big) \\ = & co\{\, J^T\lambda : J \in \mathcal{J}(x), \lambda \in \mathbb{R}^m\,\}, \end{aligned}$$

because the last set is a subspace in a finite dimensional linear space, which is closed.

The second equality in (7.5.5) is obvious. Now let $\lambda \in \mathbb{R}^m$ and $J \in \mathcal{J}(x)$. By (7.5.4), we can assume that there exists a sequence $\{x_i\} \subset C$ such that $x_i \notin \Omega_h$ (the set of nondifferentiable points of h) for $i = 1, 2, \cdots$, $x_i \to x$ and $Jh(x_i)^T\lambda \to J^T\lambda$. By the assumption and *Lemma 7.5.2*, $\mathcal{J}(x)$ contains only full rank matrices, so $Jh(x_i)$ is of full rank for i large enough. Therefore by *Theorem 7.4.3*, $Jh(x_i)^T\lambda \in \widehat{N}_C(x_i)$, So $J^T\lambda \in N_C(x)$, by the limiting property of the strict normal cone. Therefore the set of the last expression in (7.5.5) is a subset of $N_C(x)$ because of the fact that $N_C(x)$ is a closed convex cone in finite dimensional space. Combining with (7.5.6), we finish the proof of the theorem. □

The special case for $m = 1$ in *Theorem 7.5.3*, i.e., the situation of one equality constraint, is very interesting as it provides a variation to the result we

already have about this special case in *Theorem 7.2.4*, where the condition "strict differentiability" is needed about the constraint function in order to have equality.

COROLLARY 7.5.4. *Suppose* $m = 1$ *and* h *is Lipschitzian around* x . *If* $0 \notin \partial h(x)$ *and* $\mathcal{J}_C(x) = \mathcal{J}(x)$, *then*

$$N_C(x) = co\{\, \lambda v : v \in \partial h(x), \lambda \in \mathbb{R} \,\} = span\{\partial h(x)\}.$$

Remark: The notion of the *relative limiting gradient (Jacobian)* is motivated by the *relative generalized gradient*, which is used to obtain a necessary optimality condition for $\mathcal{P}(p, q)$; see [Clar83].

Generalizations of *Theorem 7.5.3* to sets determined by both equality and inequality constraints are interesting and can be done similarly.

CHAPTER 8

CALMNESS AND THE GENERALIZED JACOBIAN

Because calmness is such an important notion in nonsmooth analysis and optimization theory as we stated in **Chapter 6**, we'll develop some sufficient conditions for calmness in this chapter. First, a straightforward sufficient condition for calmness is given, then we specify it in terms of an infimal limit with respect to a particular limit subscript set. Finally a more specific sufficient condition in terms of the generalized Jacobian and generalized gradients of the equality constraint function is obtained, and a number of applications are presented under more favorable conditions.

8.1 Calmness and the Normal Cone to the Constraint Set

A counterexample in §6.3 shows that the assertion *almost all problems are calm* is not true for Lipschitz optimization problems that incorporate equality constraints and an abstract constraint. However we don't know whether the analogous assertion is true or not for Lipschitz optimization problems that incorporate only equality constraints.

We consider this kind of optimization problem in this chapter, which is denoted by $\mathcal{P}(q)$:

$$\begin{cases} \min f(x) \\ s.t.\ \ h_j(x) = q_j, \quad j = 1, 2, \ldots, m, \end{cases} \tag{8.1.1}$$

where $f, h_j : \mathbb{R}^n \to \mathbb{R}$ are Lipschitz continuous for $j = 1, \ldots, m$ unless otherwise stated. The corresponding value function is denoted by $V(q)$. For any

$q \in \mathbb{R}^m$, in the range of $h := (h_1, \cdots, h_m)$, denote

$$C_q = \{ x \in \mathbb{R}^n : h(x) = q \}.$$

The following lemma, for which a much simpler proof is provided, is a special case of *Lemma 2.1* from [Sach88], where a set-valued constraint in Banach space was considered.

LEMMA 8.1.1. *Let* $f(x)$ *be locally Lipschitzian and* $h_j(x)$ *be continuous for* $j = 1, \ldots, m$. *Suppose* x_0 *is a local minimizer of* $\mathcal{P}(q)$, *then the problem is calm at* x_0 *if*

$$\liminf_{\substack{y \to x_0 \\ y \notin C_q}} \frac{|h(y) - h(x_0)|}{dist(y, C_q)} > 0. \tag{8.1.2}$$

PROOF: Because $f(x)$ is locally Lipschitzian, there exist ϵ, K both positive such that whenever $x, y \in x_0 + \epsilon\mathcal{B}$, we have

$$|f(y) - f(x)| \leq K|y - x|. \tag{8.1.3}$$

By (8.1.2), there exist $\epsilon', \delta > 0$ s.t. whenever $|y - x_0| \leq \epsilon'$,

$$|h(y) - h(x_0)| \geq \delta \cdot dist(y, C_q). \tag{8.1.4}$$

Let $\epsilon_0 = \min\{ \frac{\epsilon}{2}, \frac{\epsilon'}{2} \}$. As h is continuous, and C_q is a closed set, then for all $y \in x_0 + \epsilon_0\mathcal{B}$, there exists $x' \in C_q$ s.t. $dist(y, C_q) = dist(y, x')$. Therefore

$$\begin{aligned} |x' - x_0| &\leq |x' - y| + |y - x_0| \\ &\leq 2|y - x_0| = 2\epsilon_0 = \min\{ \epsilon', \epsilon \}. \end{aligned} \tag{8.1.5}$$

As x_0 is a local minimum, and by (8.1.3) and (8.1.4),

$$\begin{aligned} f(y) - f(x_0) &\geq f(y) - f(x') \\ &\geq -K|y - x'| = -K \cdot dist(y, C_q) \\ &\geq -\frac{K}{\delta}|h(y) - h(x_0)|, \end{aligned} \tag{8.1.6}$$

which proves that $\mathcal{P}(q)$ is calm at x_0. □

As we are interested in Lipschitz optimization problems that incorporate explicit equality constraints, we specify the above sufficient condition in terms of the generalized gradient and the proximal normal cone.

We recall that the proximal normal cone is not necessarily closed, and, from [RoWe93], the closed convex cone of the limiting proximal normals is the Clarke normal cone, which is the cone generated by the closure of the generalized gradient of the distance function $d_{C_q}(x)$, i.e.

$$cl\big(co\{\limsup_{\substack{x\to x_0\\ x\in C_q}} PN_{C_q}(x)\}\big) = N_{C_q}(x_0), \tag{8.1.7}$$

and

$$N_{C_q}(x) = cl\{\cup_{\lambda\geq 0}\lambda\partial d_{C_q}(x)\}, \tag{8.1.8}$$

Formula (8.1.7) is referred to as the proximal normal formula, and has been extended to Hilbert spaces, see [Loew85].

Denote

$$S(x) := \{(\{x_i\}, v) : x_i \in C_q : x_i \to x, \exists\, v_i \in PN_{C_q}(x_i) \quad \text{s.t.} \quad v_i \to v\}, \tag{8.1.9}$$

then we can specify *Lemma 8.1.1* in the following way:

THEOREM 8.1.2. *Under the same conditions on* $f(x)$, $h(x)$ *and* x_0 *as in Lemma 8.1.1. If we have*

$$\inf_{(\{x_i\},v_0)\in S(x_0)} \liminf_{\substack{t\downarrow 0\\ i\to\infty}} \frac{|h(x_i+tv_0)-h(x_i)|}{t} > 0, \tag{8.1.10}$$

then problem $\mathcal{P}(q)$ *is calm at* x_0.

PROOF: It suffices to show that (8.1.2) holds. If (8.1.2) is not true, then there exists $\{y_i\}$ s.t. $y_i \notin C_q$, $y_i \to x_0$ and $|h(y_i)-h(x_0)|/dist(y_i, C_q) \to 0$.

For given i, there exists $x_i \in C_q$ such that $dist(y_i, C_q) = |y_i - x_i|$. Let $t_i = |y_i - x_i|$. Then $x_i \to x_0$ and $t_i \downarrow 0$. Let $v_i = (y_i - x_i)/t_i$. Then $|v_i| = 1$ and v_i is a proximal normal to C_q at x_i, i.e. $v_i \in PN_{C_q}(x_i) \subset N_{C_q}(x_i)$. The sequence $\{v_i\}$ has a convergent subsequence. Without loss of generality, suppose $v_i \to v_0$, then $(\{x_i\}, v_0) \in S(x_0)$. Therefore

$$\begin{aligned}
&\liminf_{i\to\infty} \frac{|h(y_i) - h(x_0)|}{dist(y_i, C_q)} = \liminf_{i\to\infty} \frac{|h(x_i + t_i v_i) - h(x_i)|}{t_i} \\
&\geq \liminf_{i\to\infty} \frac{|h(x_i + t_i v_0) - h(x_i)|}{t_i} - \limsup_{i\to\infty} \frac{|h(x_i + t_i v_0) - h(x_i + t_i v_i)|}{t_i} \\
&\geq \liminf_{i\to\infty} \frac{|h(x_i + t_i v_0) - h(x_i)|}{t_i} - \limsup_{i\to\infty} K'|v_i - v_0| \\
&= \liminf_{i\to\infty} \frac{|h(x_i + t_i v_0) - h(x_i)|}{t_i} > 0,
\end{aligned}$$

where K' is the Lipschitz modulus of h. This is a contradiction and hence the proof is completed. □

Notice that by (8.1.9), $v_i \in PN_{C_q}(x_i)$. Thus by using the relationship in §7.4, we can specify further the sufficient condition in the above theorem.

8.2 Calmness under more Favorable Conditions

As we mentioned in §6.1, calmness is the weakest constraint qualification for Lipschitz optimization problems. Considering problem $\mathcal{P}(q)$, we can specify the corresponding *Mangasarian-Fromowitz condition* at a point x in the nonsmooth case:

There is no $v_i \in \partial h_i(x)$ for $i = 1, \cdots, m$ s.t. $v_1, \cdots, v_m$ are linearly dependent.

Applied to $\mathcal{P}(q)$, the Lagrange multiplier theorem says: if x_0 is a local minimum of $\mathcal{P}(q)$, then $0 \in \partial_x(\lambda f(x_0) + \langle v, h(x_0)\rangle)$, therefore $0 \in \lambda \partial f(x_0) +$

$\sum_{i=1}^{m} v_i \partial h_i(x_0)$ for some $\lambda \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$ not all zero. Under the *Mangasarian-Fromowitz condition*, if $\lambda = 0$, then $v = 0$. By *Corollary 5* in section 6.5 from [Clar83], the problem $\mathcal{P}(q)$ is calm at x_0. So the *Mangasarian-Fromowitz condition* implies calmness.

Due to *Corollary 7.4.4*, we have the following sufficient condition for calmness.

THEOREM 8.2.1. *Consider $\mathcal{P}(q)$ defined by (8.1.1) for q in the range of Lipschitz function h. Let x_0 be a local minimizer of $\mathcal{P}(q)$, $f(x)$ is locally Lipschitzian at x_0 and suppose*

(i) h is differentiable around x_0 ;

(ii) for any given two limiting Jacobian of h at x_0 J and J', i.e. $J, J' \in \mathcal{J}(x_0)$, $J'J^T$ as an $n \times n$ matrix is of full rank.

Then $\mathcal{P}(q)$ is calm at x_0.

PROOF: If (8.1.10) is not true, then by the nature of the subscript set of infimal in (8.1.10), there exists $(\{x_i\}, v_0) \in S(x_0)$, defined by (8.1.9), such that

$$\liminf_{\substack{t \downarrow 0 \\ i \to \infty}} \frac{|h(x_i + tv_i) - h(x_i)|}{t} = 0, \tag{8.2.1}$$

for some $v_i \in PN_C(x_i)$, i.e. the infimal can be reached by an element in $S(x_0)$. Then there exists a sequence $\{t_i\}$ s.t. $t_i \downarrow 0$ and

$$|h(x_i + t_i v_i) - h(x_i)|/t_i \to 0. \tag{8.2.2}$$

By (i), x_i is eventually differentiable. Because $v_i \in PN_C(x_i)$, from *Corollary 7.4.4*, $v_i = Jh(x_i)^T \lambda_i$ for some $\lambda_i \in \mathbb{R}^m$. Without loss of generality, we assume $|v_i| = 1$, $Jh(x_i) \to J$ and $\lambda_i \to \lambda_0$. $|\lambda_0| \neq 0$ as $|v_i| \to |v_0| = 1$.

Therefore, by the mean value theorem,

$$\begin{aligned} &\lim_{i\to\infty} |h(x_i + t_i v_i) - h(x_i)|/t_i \\ &= \lim_{i\to\infty} Jh(x_i')v_i = \lim_{i\to\infty} Jh(x_i')Jh(x_i)^T\lambda_i \\ &= \lim_{i\to\infty} J'J^T\lambda_0, \end{aligned} \tag{8.2.3}$$

for some $x_i' \in (x_i, x_i + t_i v_i)$, which can not be 0 because $\lambda_0 \neq 0$ and $J'J^T$ is a full rank matrix by (ii). This is a contradiction to (8.2.2) and the theorem is proved. □

The conditions of the above theorem are not stronger than the *Mangasarian-Fromowitz condition*, because condition (ii) doesn't mean that $\partial h(x)$ contains only full rank matrices, which is implied by the *Mangasarian-Fromowitz condition*. There are functions differentiable everywhere, but not continuously differentiable. The conditions of the above theorem are also not weaker than the *Mangasarian-Fromowitz condition*, because of condition (i).

Under more favorable conditions, we have the following result.

THEOREM 8.2.2. *Suppose* $h : \mathbb{R}^n \to \mathbb{R}^m$ *is regular, differentiable at* x_0, *and Lipschitzian around* x_0. *If* x_0 *is a local minimizer of* $\mathcal{P}(q)$, *then* $\mathcal{P}(q)$ *is calm at* x_0 *under the condition that the Jacobian of* $h = (h_1, \ldots, h_m)$ *at* x_0 *is of full rank.*

PROOF: Because h is regular at x_0 and differentiable at x_0, h is strictly differentiable at x_0, from 5C.15 in [RoWe93]. So $\partial h(x_0)$ is a singleton, i.e. $\partial h(x_0)$ contains one full rank matrix. Therefore the result follows by the *Mangasarian-Fromowitz condition* specified at the beginning of this section. □

The similar problem with a set-valued constraint mapping along with an abstract constraint was considered in a Banach space in [Sach88], where a tougher

condition called *regularity of constraint*, which means the differentiability of the constraint function $h(x)$ around the local minimum, was used. *Regularity of constraint* was introduced for a single-valued mapping by Penot [Peno82], Zowe and Kurcyusz [ZoKu79] and generalized to a set-valued mapping by Dien [Dien85] to assure calmness.

The following two theorems are direct applications of *Theorem 8.2.2* and provide cases where *almost all problems are calm*. The first application requires strong smoothness of the constraint function $h(x)$.

PROPOSITION 8.2.3. *(Sard) Let h be a $\mathcal{C}^{n-m+1}$ function from an open subset A of $\mathbb{R}^n$ into a normed vector space $\mathbb{R}^m$. Then*

$$h(\{\, x \in A : rank\, Jh(x) \leq m-1 \,\}) \tag{8.2.4}$$

as a subset of $\mathbb{R}^m$ has measure 0 (assume $m \leq n$).

PROOF: See §12.1 in [Morg88]. □

THEOREM 8.2.4. *For $n \geq m$, suppose $h(x) : \mathbb{R}^n \to \mathbb{R}^m \in \mathcal{C}^{n-m+1}$. If $f(x)$ is Lipschitzian, then $\mathcal{P}(q)$ is calm almost everywhere in any open set upon which $V(q)$ is bounded.*

PROOF: As $h \in \mathcal{C}^{n-m+1}$ and $n \geq m$, h is regular everywhere. For any q in the range of h, by *Theorem 8.2.2*, $\mathcal{P}(q)$ is calm at any local minimizer x (of $\mathcal{P}(q)$) where $Jh(x)$ has full rank, i.e. $rank\, Jh(x) = m$. That is, $\mathcal{P}(q)$ is calm for all q not in the set of (8.2.4). Therefore, by *Proposition 8.2.3*, $\mathcal{P}(q)$ is calm almost everywhere for q in any open set upon which $V(q)$ is bounded. □

The second application is for a special case; it requires that the number of equality constraints equals the number of variables in the optimization problem.

PROPOSITION 8.2.5. *Suppose* $h(x) : S \in \mathbb{R}^n \to \mathbb{R}^n$ *is Lipschitzian and* S *is open, then*

$$h(\{\, x : h \text{ differentiable at } x \text{ , with Jacobian not full rank}\,\}) \tag{8.2.5}$$

as a subset of $\mathbb{R}^n$ *has measure* 0 .

PROOF: (See [Mign76]). □

THEOREM 8.2.6. *Suppose* $h : \mathbb{R}^n \to \mathbb{R}^n$ *is regular almost everywhere and Lipschitzian. If* $f(x)$ *is Lipschitzian, then* $P(q)$ *is calm almost everywhere for* q *in any open set upon which* $V(q)$ *is bounded.*

PROOF: Denote by Ω_h the set of non-differentiable points or irregular (not regular) points of h . As $h : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitzian, then Ω_h and $h(\Omega_h)$ are both sets of measure zero. For any q in the range of h , $\mathcal{P}(q)$ is calm at any local minimizer (of $\mathcal{P}(q)$) where h is differentiable and the Jacobian is of full rank. That is, for any q not in the set of (8.2.5) and $h(\Omega_h)$, $\mathcal{P}(q)$ is calm. Because $h(\Omega_h)$ and (8.2.3) are both sets of measure zero, therefore $\mathcal{P}(q)$ is calm almost everywhere for q in any open set upon which $V(q)$ is bounded. □

Notice that a special case of *Theorem 8.2.4* as $n = m$ requires $h(x) : \mathbb{R}^n \to \mathbb{R}^n \in \mathcal{C}^1$, i.e. $h(x)$ continuous differentiable. Therefore *Theorem 8.2.6* is stronger than the special case of *Theorem 8.2.4*.

8.3 Conclusions

Calmness is a very useful notion for a number of applications such as for constraint qualifications, exact penalization problems, and sensitivity analysis. Clarke showed that "*almost all problems are calm*" for inequality-constrained optimization problems. He raised the question of whether the similar assertion is true for

optimization problems that incorporate equality constraints and an abstract constraint. In **Chapter 6**, we provide counterexamples for the open problem, and in **Chapter 8**, we develop some sufficient conditions for calmness based on the relationships between the normal cone of equality constraint sets and the generalized gradients of the constraint functions developed in **Chapter 7**. These relationships enrich the theory of generalized gradients and various kind of normal cones and provide explicit expressions for normal cones of a manifold in finite dimensional space. However it is still open whether the question of "almost everywhere calmness" is true for Lipschitz optimization problems that incorporate only equality constraints (without an abstract constraint).

APPENDIX

INTERMEDIATE VALUE THEOREM AND THE BROUWER DEGREE

In this section we prove the intermediate value theorem in the general case, which is used to prove *Theorem 7.4.1*, by using the Brouwer degree.

Let $\Omega \subset \mathbb{R}^n$ be an open bounded subset and $f : \overline{\Omega} \to \mathbb{R}^n$ be a continuous function such that $f^{-1}(0) \cap \partial\Omega = \emptyset$, where $\partial\Omega$ and $\overline{\Omega}$ denote the boundary and the closure of Ω. The topological degree of Brouwer for f and Ω is a real number denoted by $\deg(f, \Omega)$, which satisfies the following properties:

(i) (Normalization) $\deg(f, \Omega) = 1$, if $f : \Omega \to \mathbb{R}^n$ is the identity map;

(ii) (Homotopy Invariance) $\deg(H(t, \cdot), \Omega)$ is independent of t, whenever $H : [0,1] \times \Omega \to \mathbb{R}^n$ is a homotopy and $H^{-1}(t, \cdot)(0) \cap \partial\Omega = \emptyset$ for every $t \in [0,1]$;

(iii) (Existence) If $\deg(f, \Omega) \neq 0$, then there exists $x \in \Omega$ such that $f(x) = 0$.

For more details of the Brouwer degree, we refer to [Deim85].

Let $\Omega = \{ x = (x_1, \cdots, x_n) \in \mathbb{R}^n : |x_i| < 1,\ i = 1, \cdots, n \}$, and $f := (f_1, \cdots, f_n) : \overline{\Omega} \to \mathbb{R}^n$ be continuous. We are interested in zeros of f in Ω, i.e. $x \in \Omega$ such that $f_i(x) = 0$ for $i = 1. \cdots, m$. We denote by

$$\begin{aligned} \partial_i^+\Omega &:= \{ x \in \partial\Omega : x_i = 1 \ \}; \\ \partial_i^-\Omega &:= \{ x \in \partial\Omega : x_i = -1 \}, \end{aligned} \tag{A.1}$$

which are two of the opposite sides of the n-dimensional box Ω.

LEMMA A.1. *Assume that $f : \overline{\Omega} \to \mathbb{R}^n$ is continuous, and for each $i = 1, 2, \cdots, n$, $f_i(x) > 0$ for $x \in \partial_i^+\Omega$ and $f_i(x) < 0$ for $x \in \partial_i^-\Omega$. Then there exists $\xi \in \Omega$ such that $f(\xi) = 0$.*

PROOF: By the assumption, $0 \notin f(\partial\Omega)$, so the Brouwer degree $\deg(f,\Omega)$ is well defined. Define a homotopy $H : [0,1] \times \overline{\Omega} \to \mathbb{R}^n$ as follows:

$$H(t,x) := (1-t)f(x) + tx, \quad (t,x) \in [0,1] \times \overline{\Omega}.$$

We conclude, for any $t \in [0,1]$, $0 \notin H(t,\partial\Omega)$. This is true for $t = 1$ since $H(1,\cdot)$ is an identity map. If this is not true for some $t' \in [0,1)$, then there exists $x \in \partial\Omega$ with $x_i = 1$ or -1 for some integer $i \in \{1,2,\cdots,m\}$ such that $H(t',x) = 0$, i.e. $x \in \partial_i^+\Omega$ or $x \in \partial_i^-\Omega$. This implies $f_i(x) = -t'x_i/(1-t') \le 0$ for $x \in \partial_i^+\Omega$ or $f_i(x) = -t'x_i/(1-t') \ge 0$ for $x \in \partial_i^-\Omega$, which contradicts the assumption. Therefore $\deg(H(t,\cdot),\Omega)$ is well defined for each $t \in [0,1]$ and its value is independent of $t \in [0,1]$. Note that $H(1,\cdot)$ is an identity map. By (i) and (ii), we have

$$\deg(f,\Omega) = \deg(f(H(0,\cdot),\Omega) = \deg(H(1,\cdot),\Omega) = 1.$$

Now the conclusion follows from the existence property (iii) of the degree, and the proof is completed. □

The above lemma can be easily extended to a more general intermediate value theorem. Consider an open set

$$\Omega' := \{\, x \in \mathbb{R}^n : \alpha_i < x_i < \beta_i,\ i = 1,2,\cdots,n \,\},$$

for $\alpha := (\alpha_1,\cdots,\alpha_n), \beta := (\beta_1,\cdots,\beta_n) :\in \mathbb{R}^n$ such that $\alpha_i < \beta_i$ for $i = 1,2,\cdots,n$. We define $\partial_i^+\Omega'$ and $\partial_i^-\Omega'$ in the same way as (A.1).

THEOREM A.2. *If f is a continuous function defined on an open set containing $\overline{\Omega'}$, and for each $i = 1,2,\cdots,n$, $f_i(x) > a_i$ for $x \in \partial_i^+\Omega'$ and $f_i(x) < a_i$ for $x \in \partial_i^-\Omega'$. Then there exists $\xi \in \Omega'$ such that $f(\xi) = a := (a_1,a_2,\cdots,a_n)$.*

PROOF: Under a linear transformation of variables and a shift of functions, this

is equivalent to Lemma A.1. □

In the case of $m = 2$, there is a stronger result in [MuWi74].

REFERENCES

[Armi66] L. Armijo, "Minimization of functions having Lipschitz-continuous first partial derivatives", *Pacific J. Math.*, **16**, pp. 1-3 (1966).

[AuFr90] J.-P. Aubin and H. Frankowska, **Set-Valued Analysis**, (Birkhäuser, Boston) (1990).

[Avri76] M. Avriel, **Nonlinear Programming, Analysis and Methods**, (Prentice-Hall, New Jersey) (1976).

[Axel74] O. Axelsson, " On preconditioning and convergence acceleration in sparse matrix problems", *Report 74-10, CERN European Organization for nonlinear research, Geneva*, (1974).

[BDMo73] C.G. Broyden, J.E. Jr. Dennis, J.J. Moré, "On the local and superlinear convergence of quasi-Newton methods", *J. Inst. Maths. Applics.*, **12**, pp.223–245 (1973).

[Broy70] C.G. Broyden, "The convergence of a class of double-rank minimization algorithms", *J. Inst. Maths. Applics.*, **6**, pp. 76–90 (1970).

[Burk91] J.V. Burke, "Calmness and exact penalization", *SIAM J. Control Optim.*, **29**, pp. 493-497 (1991).

[BySS87] R.H. Byrd, R.B. Schnabel and G.A. Shultz, "A trust region algorithm for nonlinearly constrained optimization", *SIAM J. Numer. Anal.*, **24**, pp.1152–1170 (1987).

[Carr61] C.W. Carroll, "The created response surface technique for optimizing nonlinear restrained system", *Operations Research*, **9**, pp.169-184 (1861).

[Cauc1847] A. Cauchy, "Méthode Générale pour la Résolution des Système d'Équations Simultanées", *Comp. Rend. Acad. Sci. Paris*, pp.536–538 (1847).

[CGOL76] P. Concus, G.H. Golub and D.P. O'Leary, "A generalized conjugate-gradient method for the numerical solution of elliptic partial differential equations", **Sparse Matrix Computations, J.R. Bunch and D.J. Rose, eds.**, (Academic Press, London and New York) pp.309-332 (1976).

[Char73] C. Charalambous, "Unconstrained optimization based on homogeneous models", *Math. Programming*, **5**, pp.189–198 (1973).

[Clar76] F.H. Clarke, "A new approach to Lagrange multipliers", *Math. Oper. Res.*, **1**, pp.165–174 (1976).

[Clar83] F.H. Clarke, **Optimization and Nonsmooth Analysis**, (John Willey and Sons) (1983).

[Clar86] F.H. Clarke, " Perturbed Optimal control problems", *IEEE Transactions on Automatic Control*, **AC-31, No. 6**, (June, 1986).

[CoGT88] A.R. Conn, N.I.M. Gould and Ph. L. Toint, " Global convergence of a class of trust region algorithms for optimization with simple bounds", *SIAM J. Numer. Anal.*, **25**, pp. 433–460 (1988).

[Cour62] R. Courant, **Calculus of Variations and Supplementary Notes and Exercises**, ((mimeographed notes), supplementary notes by M. Kruskal and H. Robin, Revised and amended by J. Moser, New York University) (1962).

[Davi59] W.C. Davidon, "Variable metric methods for minimization", *A. E. C. Res. and Develop. Report ANL-5990, Argonne National Laboratory, Argonne, Illinois*, (1959).

[Davi80] W.C. Davidon, "Conic approximation and collinear scaling for optimizers", *SIAM J. Numer. Anal.*, **17**, pp.268–281 (1980).

[Deim85] K. Deimling, **Nonlinear Functional Analysis** , (Springer-Verlag) (1985).

[DeMe79] J.E. Dennis Jr. and M.H.W. Mei, " Two new unconstrained optimization algorithms which use function and gradient values", *J. Optim. Theory Appl.*, **28**, pp.453–482 (1979).

[DeMo74] J.E. Dennis Jr. and J.J. Moré, " A characterization of superlinear convergence and its application to quasi-Newton methods", *Math. of Compu.*, **28**, pp. 549–560 (1974).

[DeMo77] J.E. Dennis Jr. and J.J. Moré, " Quasi-Newton methods, motivation and theory", *SIAM Review*, **19**, pp. 46–89 (1977).

[DeSc83] J.E. Dennis, Jr. and R.B. Schnabel, **Numerical methods for unconstrained optimization and nonlinear equations**, (Prentice-Hall, New Jersey) (1983).

[Dien85] Pham Huy Dien, " On the regularity condition for the extremal problem under locally Lipschitz inclusion constraints", *Appl. Math. Optim.*, **13**, pp.151-161 (1985).

[Ermi66] I.I. Eremin, " The 'penalty' method in convex programming", *Dokl. Akad. Nauk SSSR (English)*, **173**, pp.459-462 (1966).

[Fiao83] A.V. Fiacco, **Introduction in Sensitivity and Stability Analysis in Nonlinear Programming**, (Academic Press, New York) (1983).

[FiMc68] A.V. Fiacco and G.P. McCormick, **Nonlinear Programming: Sequential Unconstrained Minimization Techniques**, (John Wiley and Sons, New York and Toronto) (1968).

[Flet70] R. Fletcher, " A new approach to variable metric algorithms", *Computer Journal*, **13**, pp.313–322 (1970).

[Flet71] R. Fletcher, " A modified Marquardt subroutine for nonlinear least squares", *Report R6799, Atomic Energy Research Establishment, Harwell, England*, (1971).

[Flet72] R. Fletcher, " An algorithm for solving linearly constrained optimization problems", *Math. Prog.*, **2**, pp.133-195 (1972).

[Flet80] R. Fletcher, **Practical Methods of Optimization, V. 1: Unconstrained optimization**, (John Wiley & Sons) (1980).

[FlPo63] R. Fletcher, M.J.D. Powell, " A rapid convergence descent method for minimization", *Computer Journal*, **6**, pp.163–168 (1963).

[FlPo64] R. Fletcher, C.M. Reeves, " Function minimization by conjugate gradients", *Computer Journal*, **7**, pp.149–154 (1964).

[Fric71] I. Fried, " N-step conjugate gradient minimization scheme for non-quadratic functions", *A.I.A.A.Journal*, **7**, pp.2286–2287 (1971).

[GaDu84] J. Gauvin and F. Dubeau, " Some examples and counterexamples for the stability analysis of nonlinear programming problems", *Math. Programming Stud.*, **21**, pp.69–78 (1984).

[Gay79] D.M. Gay, " On robust and generalized linear regression problems", *Report 2000, Mathematics Research Center, University of Wisconsin, Madison, Wisconsin*, (1979).

[Gay81] D.M. Gay, " Computing optimal local constrained step", *SIAM J. Sci. Stat. Comput.*, **2**, pp.186–197 (1981).

[Gay83] D.M. Gay, " Algorithm 611, subroutines for unconstrained minimization using a model/ trust-region approach", *ACM Trans. Math. Softw.*, **9**, pp. 503–524 (1983).

[GiMu72] P.E. Gill and W. Murray, " Quasi-Newton methods for unconstrained optimization", *J. Inst. Math. Appl.*, **9**, pp.91–108. (1972).

[GMWr81] P.E. Gill, W. Murray and M.H. Wright, **Practical Optimization**, (Academic Press) (1981).

[Gold67] A.A. Goldstein, **Constructive real analysis**, (Harper & Row, New York) (1976).

[Gold70] D. Goldfarb, "A family of variable metric methods derived by variational means", *Mathematics of Computation*, **24**, pp.23–26 (1970).

[GQTr66] S.M. Goldfeld, R.E. Quandt and H.F. Trotter, "Maximization by quadratic hill-climbing", *Econometrica*, **34**, pp.541–551 (1966).

[Hebd73] M.D. Hebden, " An algorithm for minimization using exact second derivatives", *Rept. TP515, A.E.R.E. Harwell, England*, (1973).

[HeSt52] M.R. Hestenes and E.L. Stiefel, **Methods of conjugate gradients for solving linear system**, *J. Res. Nat. Bur. Standards, Section B*, **49**, pp.409–436 (1952).

[Hest80] M.R. Hestenes, **Conjugate Gradient Methods in Optimization**, (Springer–Verlag, New York) (1980).

[Hira79] J.B. Hiriart-Urruty, " Tangent cones, generalized gradients and mathematical programming in Banach spaces", *Math. Oper. Res.*, **4**, pp.79–97 (1979).

[IMSL10] **IMSL Manual, Version 10**, (1988).

[JaOx72] D. Jacobson and W. Oxman, " An algorithm that minimizes homogeneous functions of n variables in $N+2$ iterations and rapidly minimizes general functions", *J. Math. Anal. Appl.*, **38**, pp.533–552 (1972).

[KaAk64] L.V. Kantorovich and G.P. Akilov, **Functional Analysis in Normed Space**, (Macmillan, New York) (1964).

[Kiw85a] K.C. Kiwiel, " Methods of descent for nondifferentiable optimization", **Lecture Notes in Mathematics 1133**, (Springer–Verlag, Berlin) (1985).

[Kiw85b] K.C. Kiwiel, " A linearization algorithm for nonsmooth minimization", *Math. Oper. Res.*, **10**, pp.185–194 (1985).

[Kiw89a] K.C. Kiwiel, " A Dual method for solving certain positive semidefinite quadratic programming problems", *SIAM J. Sci. Stat. Comput.*, **10**, pp.175–186 (1989).

[Kiw89b] K.C. Kiwiel, " An ellipsoid trust region bundle method for nonsmooth convex minimization", *SIAM J. Control. Optim.*, **27**, pp.737–757 (1989).

[Kiw90] K.C. Kiwiel, " Proximity control in bundle methods for convex non-differentiable minimization", *Math. Programming*, **46**, pp.105–122 (1990).

[Leve44] K. Levenberg, " A method for solving certain problems in least squares", *Quart. Appl. Math.*, **2**, pp.164-168 (1944).

[Loew85] P.H. Loewen, " Proximal normal analysis and dynamic optimization", *Ph.D. dissertation, U.B.C.* , (1985).

[Luen73] D.G. Luenberger, **Introduction to Linear and Nonlinear Programming**, (Addison-Wesley, Menlo Park, California) (1973).

[Marq63] D. Marquardt, " A algorithm for least-squares estimation of nonlinear parameters", *SIAM J. Appl. Math.*, **11**, pp.431–441 (1963).

[Mign76] F. Mignot, "Contrôle dans les inéquations variationelles elliptiques", *J. of Func. Anal*, **22**, pp.pp130-185 (1976).

[MiPe92] J.P. Michel and J.-P. Penot, "A generalized derivative for calm and stable functions", *Differential and Integral Equations*, **5**, pp.433–454 (1992).

[MoGH81] J.J. Moré, B.S. Garbow and K.E. Hillstorm, "Testing unconstrained optimization software", *ACM Trans. Math. Softw.*, **7**, pp.17-41 (1981).

[More77] J.J. Moré, " The Levenberg-Marquardt algorithm: implementation and theory", **Numerical Analysis, G.A. Watson, ed. Lecture Notes in Math. 630, Springer Verlag, Berlin**, pp.105–116. (1977).

[More82] J.J. Moré, " Recent development in algorithms and software for trust region methods", **Mathematical Programming, The State of the Art, Bonn**, A. Bachem, M. Grötschel, B. Korte ed. pp.258–287. (1982).

[Morg88] F. Morgan, **Geometric measure theory**, (Academic Press) (1988).

[MoSo83] J.J. Moré and D.C. Sorensen, "Computing a trust region step", *SIAM J. Sci. Stat. Comput.*, **4**, pp.553–572 (1983).

[MuWi74] J.S. Muldowney and D. Willett, "An elementary proof of the existence of solutions to second order nonlinear boundary value problems", *SIAM J. Math. Anal.*, **5**, pp.701–707 (1974).

[Naza79] L. Nazareth, " Relationship between the BFGS and conjugate-gradient algorithms and its implications for new algorithms", *SIAM J. Numer. Anal.*, **16**, pp.794–800 (1979).

[Peno82] J.P. Penot, "On regularity conditions in mathematical programming", *Math. Programming Stu.*, **19**, pp.167-199 (1982).

[Piet69] T. Pietrzykowski, " An exact potential method for constraint maxima", *SIAM J. Numer. Anal.* , **6**, pp. 299-304 (1969).

[Poli88] R. Poliquin, " Proto-differentiation and Integration of Proximal Subgradients", *Ph.D. Dissertation, University of Washington*, (1988).

[Powe70] M.J.D. Powell, " A hybrid method for nonlinear equations", **Numerical Methods for Nonlinear Algebraic Equations**, pp. 87-114 (p. Robinowitz, ed., Gordon and Breach, London) (1970).

[Powe71] M.J.D. Powell, " On the convergence of variable metric algorithm", *J. Inst. Maths. Applics.*, **7**, pp. 21-36 (1971).

[Powe75] M.J.D. Powell, " Convergence properties of a class of minimization algorithms", **Nonlinear Programming 2 (O.L. Mangasarian, R.R. Meyer and S.M. Robinson, eds.**, pp. 1-27 (Academic Press, London and New York) (1975).

[Rock70] R.T. Rockafellar, **Convex Analysis**, (Princeton, University Press) (1970).

[Rock81] R.T. Rockafellar, "Proximal subgradients, marginal values, and augmented Lagrangian in nonconvex optimization", *Math. Oper. Res.*, **6**, pp. 427–437 (1981).

[Rock85] R.T. Rockafellar, "Extensions of subgradient calculus with applications to optimization", *Nonlinear Anal.*, **9**, pp. 665-698 (1985).

[RoWe93] R.T. Rockafellar and R. J.-B. Wets, **Variational Analysis**, to appear, (Springer-Verlag)

[Sach88] Pham Huu Sach, " Regularity, Calmness and Support Principle", *Optimization*, **19**, pp.13-27 (1988).

[Saks33] S. Saks, **Théorie de L'intégrale**, (Z Subwencji Funduszu Kultury Narodowej Warszawa) (1933).

[Schn77] R.B. Schnabel, "Analyzing and improving quasi-Newton methods for unconstrained optimization", (Ph.d. Thesis, TR77-320, Dept. of Comp. Sci., Conell Univ., NY, 1977.)

[ScZo88] H. Schrammm and J. Zowe, "A combination of bundle approach and the trust region concept", **Advances in Mathematical Optimization**, (J. Guddat et al., eds., Akademie–Verlag, Berlin) pp.196–209 (1988).

[ScZo90] H. Schrammm and J. Zowe, "A version of bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results", *Report No. 206, Mathematisches Institut, Universität Bayreuth*, (1990).

[Shan70] D.F. Shanno, "Conditioning of quasi-Newton methods for function minimization", *Mathematics of Computation*, **24**, pp.647–657 (1970).

[ShSB85] G.A. Shultz, R.B. Schnabel and R.H. Byrd, "A family of trust-region-based algorithms for unconstrained minimization with strong global convergence properties", *SIAM J. Numer. Anal.*, **22**, pp.47–66 (1985).

[Sore80] D.C. Sorensen, "The Q-superlinear convergence of a collinear scaling algorithm for unconstrained optimization", *SIAM J. Numer. Anal.*, **17**, pp.84–114 (1980).

[Sore80a] D.C. Sorensen, " Newton's method with model trust region modification", *Report ANL-80-106, Argonne National Laboratory, Argonne, Illinois*, (1980).

[Sore82a] D.C. Sorensen, "Trust region method for unconstrained minimization", **Nonlinear Optimization 1981, M.J.D. Powell ed.**, pp.29-46 (1982).

[Sore82b] D.C. Sorensen, "Newton's method with a model trust region modification", *SIAM J. Numer. Anal.*, **19**, pp. 409–426 (1982).

[Stoe75] J. Stoer, " On the convergence rate inperfect minimization algorithms in Broyden's β-class", *Math. Prog.*, **9**, pp.313–335 (1975).

[Ward90] D. Ward, "Calmness, exact penalties and constraint qualifications in non-Lipschitzian optimization", *Preprint*, (1990).

[WoJi90] Y.S. Wong and H. Jiang, "Approximate polynomial preconditioning applied to biharmonic equations", *J. Supercomputing*, **8**, pp. 226–235 (1990).

[Wolf69] P. Wolfe, "Convergence conditions for ascent methods", *SIAM Review*, **11**, pp. 226–235 (1969).

[Wolf71] P. Wolfe, "Convergence conditions for ascent methods. I: some corrections", *SIAM Review*, **13**, pp. 185–188 (1971).

[Zang67] W.I. Zangwill, " Nonlinear Programming via penalty functions", *Management Science*, **13**, pp.344-358 (1967).

[ZoKu79] J. Zowe and S. Kurcyusz, "Regularity and Stability for mathematical programming in Banach spaces", *Appl. Math. Optim.*, **5**, pp.49-62 (1979).

[Zowe85] J. Zowe, "Nondifferentiable optimization", **Computational Mathematical Programming**, (K. Schittkowski ed., Springer–Verlag, Berlin) pp.323-356 (1985).